

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 16:59:45 ; Search time 75 Seconds

(without alignments)
365.995 Million cell updates/sec

Title: US-09-887-855-5

Perfect score: 1115
Sequence: 1 ATGRLLSGQPVCRGQTQRPC.....EEDAKTFKESREAAALNAY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: A_Geneseq_101002:*
 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
23: 
```

1115 100.0 374 21 AAY93948

1115 100.0 374 22 AAE03651

1115 100.0 374 23 ABG66680

1115 100.0 374 23 ABB90203

1105 99.1 374 22 AAM25796

1101 98.7 382 21 AAY91490

1101 98.7 382 20 AAY13367

1101 98.7 382 22 AAU29033

1101 98.7 382 22 AAB80235

561.5 50.4 273 21 AAB18913

11 561.5 50.4 273 22 AAU12441
12 561.5 50.4 273 22 AAB73309
13 561.5 50.4 273 22 AAB87609
14 561.5 50.4 273 23 ABB95554
15 561.5 50.4 273 23 ABB84948
Human protein sequ
Protein #3787 enco
Peptide #3978 enco
Human secreted pro
Human secreted pro
Human secreted pro
Human secreted pro
Soluble mannose re
Novel human diagno
Murine macrophage
Human novel extrac
Human polypeptide
Human type C lecti
Human ORF2915
Human type C lecti
Human Tumour Endot
Human novel extrac
Human polypeptide
Novel human diagno
Human versican iso
Human cancer assoc
Neurocan core prot
Versican. Homo sa
Murine type C lect
Murine E-selectin
Mouse Tumour Endot
Amino acid sequenc
Mouse protein enco
Human DC-SIGN prot
Mouse cell surface
Haematopoietic ste

ALIGNMENTS

```

RESULT 1
ID AAY93948 standard; Protein; 374 AA.
XX
AC AAY93948;
XX DT 03-OCT-2000 (first entry)
XX DE Amino acid sequence of a lectin ss3939 polypeptide.
XX KW Human; lectin ss3939; chromosome 11; gene therapy.
XX OS Homo sapiens.
XX
Key Location/Qualifiers
 1..21 Peptide /note= "signal peptide"
Domain Domain /note= "extracellular coding region"
 22..227 /note= "predicted transmembrane domain"
 228..248 /note= "predicted cytoplasmic or intracellular domain"
 249..374 /note= "predicted cytoplasmic or intracellular domain"

WO200039296-A1.
PN XX
XX
PD 06-JUL-2000.
XX
PF 22-DEC-1999;
XX
PR 23-DEC-1998;
XX
PA (IMMV ) IMMUNEX CORP .
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	1115	100.0	374	21	AAY93948		Amino acid sequenc
2	1115	100.0	374	22	AAE03651		Human extracellula
3	1115	100.0	374	23	ABG66680		Human novel polype
4	1115	100.0	374	23	ABB90203		Human polypeptide
5	1115	100.0	387	22	AAM25796		Human protein sequ
6	1105	99.1	374	21	AAY91490		Human secreted pro
7	1101	98.7	382	20	AAY13367		Amino acid sequenc
8	1101	98.7	382	22	AAU29033		Human PRO Polypept
9	1101	98.7	382	22	AAB80235		Human PRO234 prote
10	561.5	50.4	273	21	AAB18913		A novel polypeptid

XX Anderson DA;
 XX WPI; 2000-452394/39.
 DR N-PSDB; AAA57382.

XX ss3939 nucleic acids, polypeptides and antibodies, useful for identifying human chromosome 11 and diseases associated with it -

XX Claim 12; Page 8; 73pp; English.

XX The present sequence represents a human lectin ss3939 polypeptide. The polynucleotide sequence is a source of probes, which may be used to identify nucleic acids encoding ss3939 proteins, to identify human chromosome number 11, to map genes on human chromosome number 11, to identify diseases associated with chromosome 11, as single-stranded sense or antisense Oligonucleotides to inhibit expression of polypeptides encoded by the ss3939 gene, and for gene therapy. The ss3939 polypeptides may be useful for developing treatments for diseases (none specified) associated with defective or insufficient amounts of the polypeptides. The antibodies may be useful for detecting the presence of ss3939 polypeptides.

XX Sequence 374 AA;

Query	Match	Score	DB	Length
QY	1 ATGRLLSGQPVCRGGTQRCPCKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKL	100.08;	21;	374;
Db	22 ATGRLLSGQPVCRGGTQRCPCKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKL	100.08;	20;	3.5e-105;
QY	61 IEKFIENLLPSDGFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV	0;	0;	0;
Db	82 IEKFIENLLPSDGFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV	0;	0;	0;
QY	121 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV	180		
Db	142 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV	201		
QY	181 LPEETQEDAKKTFKESREAAALNLY	206		
Db	202 LPEETQEDAKKTFKESREAAALNLY	227		

XX Query Match 100.08; Score 1115; DB 21; Length 374;
 Best Local Similarity 100.08; Pred. No. 3.5e-105;
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX FT . Protein
 FT . Signal_peptide
 25..374
 /note- "Mature human extracellular matrix and cell adhesion molecule (XMAD)."
 46..63
 /note- "C-type lectin domain"
 163..176
 /note- "C-type lectin domain"
 224..247
 /note- "Transmembrane motif"
 328..348
 /note- "Transmembrane motif"
 WO200142285-A2.

XX Domain
 FT Domain
 FT Domain
 FT Domain
 FT Domain
 FT Domain
 PN XX
 PD XX
 PR XX
 PR XX
 PR XX
 PA (INCY-) INCYTE GENOMICS INC.

XX PI Yue H, Tang YT, Lal P, Burford N, Azimzai Y, Patterson C;
 PI Baughn MR, Lu DAM, Shah P, Au-Young J;
 WPI; 2001-381632/40.
 N-PSDB; AAD08059.

XX PT New human extracellular matrix and cell adhesion molecules and polynucleotide sequences encoding them, useful for diagnosis, prevention, treatment of genetic, autoimmune and cell proliferative disorders -

XX PT Claim 1; Page 108-109; 135pp; English.

XX The present sequence is a human extracellular matrix and cell adhesion molecule (XMAD). The XMAD is used for screening a compound for effectiveness as an agonist or antagonist of XMAD. The identified agonist or antagonist are used for treating a disease or condition associated with decreased or increased expression of functional XMAD. The polynucleotides encoding XMAD are useful in somatic or germline gene therapy to correct a genetic deficiency, to express a conditionally lethal gene product and to express a protein which affords protection against intracellular parasites and also for diagnosis of disorders associated with expression of XMAD. They are also used for generating genomic hybridisation probes useful in mapping the naturally occurring genomic sequences and to create knock in humanised animals (pigs) or transgenic animals (mice or rats) to model human diseases. Oligonucleotide or longer fragments derived from the polynucleotide sequences may be used as elements on a microarray. Antibodies which specifically bind XMAD may be used for the diagnosis of disorders associated with the expression of XMAD, or in assays to monitor patients being treated with XMAD. Diseases diagnosed, prevented or treated include genetic disorders such as adrenoleukodystrophy, Down's syndrome, cystic fibrosis, Gaucher's disease, myotonic dystrophy, sickle cell anaemia, thalassaemia, autoimmune/inflammatory disorders such as acquired immune deficiency syndrome (AIDS), Addison's disease, allergies, anaemia, asthma, atherosclerosis, Crohn's disease, diabetes mellitus, atopic dermatitis, glomerulonephritis, multiple sclerosis, Grave's disease, osteoarthritis, osteoporosis, psoriasis, rheumatoid arthritis, ulcerative colitis, bacterial, fungal, parasitic, protozoal and helminthic infections and cell proliferative disorders such as actinic keratosis, arteriosclerosis and cancer including breast, bladder, bone marrow, brain and uterus cancer, leukaemia, adenocarcinoma, lymphoma, melanoma and myeloma.

XX Sequence 374 AA;

XX Query Match 100.08; Score 1115; DB 22; Length 374;
 Best Local Similarity 100.08; Pred. No. 3.5e-105;
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 1..24

QY 1 ATGRLLSGQQVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRARDGGQLVSIESTEDEQKL 60
 CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
 CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
 CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
 CC disease. The sequences of the invention are also useful for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues, immune deficiencies and disorders
 CC including severe combined immunodeficiency (SCID), bacterial or fungal
 CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
 CC gravis, allergic conditions such as asthma, thrombolytic or thrombosis
 CC and coagulation disorders. Sequences ABG66666-ABG666758 represent human
 CC novel polypeptides of the invention.
 XX

DB 22 IEKFIENLLPSDGFDFWIGLRRREEEKQNSNTACQDLYAWTDGSISQFRNWWYDEPSCGSEV 120
 XX

QY 61 IEKFIENLLPSDGFDFWIGLRRREEEKQNSNTACQDLYAWTDGSISQFRNWWYDEPSCGSEV 141
 DB 82 IEKFIENLLPSDGFDFWIGLRRREEEKQNSNTACQDLYAWTDGSISQFRNWWYDEPSCGSEV 141

QY 121 CVVMYHQPSAPAGIGGPMQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180
 DB 142 CVVMYHQPSAPAGIGGPMQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201

QY 181 LPEETQEEAKKTFKESREALNLAY 206
 DB 202 LPEETQEEAKKTFKESREALNLAY 227

RESULT 3
 ID ABG66680 standard; Protein; 374 AA.
 XX
 AC ABG66680;
 XX
 DT 30-AUG-2002 (first entry)
 XX
 DE Human novel polypeptide #15.
 XX
 KW Human; inflammatory condition; shock; sepsis; immune response;
 KW cancer; wound healing; central nervous system disease; haematoopoiesis;
 KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;
 KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
 KW cartilage; ligament; nerve tissue; ulcer; osteoporosis;
 KW bone degenerative disorder; periodontal disease; reperfusion injury;
 KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
 KW allergic condition; thrombosis; thrombosis; coagulation disorder;
 KW fungal infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200244340-A2.
 XX
 PD 06-JUN-2002.
 XX
 PF 30-NOV-2001; 2001WO-US47004.
 XX
 PR 30-NOV-2000; 2000US-0028952.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
 PI Yamazaki V, Ujwai ML, Drmanac RT;
 DR WPI; 2002-508509/54.
 DR N-PSDB; ABK94904.
 XX
 PT Novel nucleic acids and polypeptides for diagnosis, treatment of
 PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
 PT disorders, cancer and promoting wound healing -
 XX
 PS Claim 10; Page 579-580; 672pp; English.
 XX
 CC The invention relates to human novel polynucleotides and associated
 CC polypeptides. The polynucleotides and polypeptides are useful for
 CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
 CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
 CC and cancer and for promoting wound healing. The sequences are used to
 CC induce the proliferation of neural cells and regeneration of nerve and
 CC brain tissue, and are useful for the treatment of central and peripheral
 CC nervous system diseases and neuropathies, such as Alzheimer's disease,
 CC Parkinson's disease, Huntington's disease and amyotrophic lateral
 CC sclerosis. The sequences are involved in chemotactic or chemokinetic
 CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
 CC cell disorders and platelet disorders such as thrombocytopenia,
 CC

Query Match 100.0%; Score 1115; DB 23; Length 374;
 Best Local Similarity 100.0%; Pred. No. 3.5e-105;
 Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQQVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRARDGGQLVSIESTEDEQKL 60
 Db 22 ATGRLLSGQQVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRARDGGQLVSIESTEDEQKL 81

Qy 61 IEKFIENLLPSDGFDFWIGLRRREEEKQNSNTACQDLYAWTDGSISQFRNWWYDEPSCGSEV 120
 Db 82 IEKFIENLLPSDGFDFWIGLRRREEEKQNSNTACQDLYAWTDGSISQFRNWWYDEPSCGSEV 141

Qy 121 CVVMYHQPSAPAGIGGPMQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180
 Db 142 CVVMYHQPSAPAGIGGPMQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201

Qy 181 LPEETQEEAKKTFKESREALNLAY 206
 Db 202 LPEETQEEAKKTFKESREALNLAY 227

RESULT 4
 ID ABB90203 standard; Protein; 374 AA.
 XX
 AC ABB90203;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 24-MAY-2002.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY-2001; 2001WO-US16450.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 19-MAY-2000; 2000US-205515P.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR (HUMA-) HUMAN GENOME SCI INC.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 29-NOV-2001.
 XX
 DE Human polypeptide SEQ ID NO 2579.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antibacterial; antiulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PR 18-MAY

XX	Human secreted protein sequence encoded by gene 40 SEQ ID NO:163.	DB	22 ATGRLLSGQPVCRGQTORPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSISEDEQKL 81
XX	Human; secreted protein; diagnosis; cytostatic; immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic; osteopathic; antilarthritic; antibacterial; antiasthma; antipsoriatic; cardiant; gene therapy; cancer; neurological disorder; immune disease; inflammation; blood disorder; tumour.	QY	61 IEKFIENLPSDGDFWIGLRRREEKOSNNTACQDLYAWTDGSISQFRNWTDEPSCGSEV 120
XX	Homo sapiens.	DB	82 IEKFIENLPSDGDFWIGLRRREEKOSNNTXCDLYAWTDGSISQFRNWTDEPSCGSEV 141
XX	WO200006698-A1.	QY	121 CVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180
XX	PD 10-FEB-2000.	DB	142 CVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201
XX	PF 29-JUL-1999; 99WO-US17130.	RESULT 7	
XX	PR 30-JUL-1998; 98US-0094657.	AY13367	
XX	PR 05-AUG-1998; 98US-0095486.	ID AY13367 standard; protein; 382 AA.	
XX	PR 06-AUG-1998; 98US-0095454.	XX	
XX	PR 06-AUG-1998; 98US-0095455.	AC AY13367;	
XX	PR 12-AUG-1998; 98US-0096319.	XX	
XX	PA (HUMA-) HUMAN GENOME SCI INC.	DT 25-JUN-1999 (first entry)	
XX	PI Komatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y; Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA; Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski M; WPI; 2000-195282/17.	XX	XX Amino acid sequence of protein PRO234.
XX	DR N-PSDB; AAA26385.	DE	
XX	PS Claim 11; Page 483-484; 634pp; English.	XX	XX Secreted protein; transmembrane protein; human; enterocolitis; Zollinger-Ellison syndrome; gastrointestinal ulceration; congenital microvillus atrophy; skin disease; cell growth; abnormal keratinocyte differentiation; psoriasis; epithelial cancer; Parkinson's disease; Alzheimer's disease; ALS; neuropathy; fibromodulin; dermal scarring; Usher Syndrome; Atrophy areata; anti-thrombotic; wound healing; tissue repair.
XX	PT New isolated human genes and the secreted polypeptides they encode, useful for diagnosis and treatment of e.g. cancers, neurological disorders, immune diseases, inflammation or blood disorders -	OS Homo sapiens.	
XX	PT DR N-PSDB; AAA26385.	PN WO9914328-A2.	
XX	PS	PD 25-MAR-1999.	
XX	CC The polynucleotide sequences given in AAA26346 to AAA26458 encode the human secreted proteins given in AAY91451 to AAY91691. The human secreted proteins can have activities based on the tissues and cells they are expressed in. Examples of the activities are: cytosstatic; immunosuppressive; antiHIV; antiinflammatory; nootropic; neuroprotective; antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma; antipsoriatic; and cardiant. The polynucleotides and their corresponding secreted proteins are useful for preventing, treating or ameliorating medical conditions, e.g. by protein or gene therapy. Also pathological conditions can be diagnosed by determining the amount of the proteins in a sample or by determining the presence of mutations in the polynucleotides. Specific uses are described for each of the polynucleotides, based on which tissues they are most highly expressed in, and include developing products for the diagnosis or treatment of cancer, tumours, neurodegenerative disorders, developmental abnormalities and foetal deficiencies, blood disorders, diseases of the immune system, autoimmune diseases, hepatic and renal disease, inflammation, allergies, Alzheimer's and behavioural disorders, schizophrenia, osteoporosis, arthritis, infections, AIDS, spinal cord injuries, transplant rejection, diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders, reproductive disorders, gastrointestinal disorders, respiratory disorders and metabolic disorders. The proteins or polynucleotides can also be used as food additives or preservatives. The proteins are also useful for identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are sequences used in the exemplification of the present invention.	XX	
XX	SQ Sequence 374 AA;	PR 27-OCT-1997; 97US-0063327.	
XX	SQ	PR 28-OCT-1997; 97US-0063541.	
XX	Query Match 99.1%; Score 1105; DB 21; Length 374;	PR 28-OCT-1997; 97US-0063542.	
XX	Best Local Similarity 99.0%; Pred. No. 3.6e-104;	PR 28-OCT-1997; 97US-0063549.	
XX	Matches 204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	PR 28-OCT-1997; 97US-0063550.	
XX	1 ATGRLLSGQPVCRGQTORPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSISEDEQKL 60	PR 28-OCT-1997; 97US-0063564.	
XX		PR 29-OCT-1997; 97US-0063435.	
XX		PR 29-OCT-1997; 97US-0063704.	
XX		PR 29-OCT-1997; 97US-0063732.	

PR	29-OCT-1997; PR	97US-0063738.	Db	202 ETELTPVLPETQEEDAKKTKEFRESREANLNAY 235
PR	29-OCT-1997; PR	97US-0064215.		
PR	29-OCT-1997; PR	97US-0063735.		
PR	31-OCT-1997; PR	97US-0063870.		
PR	31-OCT-1997; PR	97US-0064103.		
PR	03-NOV-1997; PR	97US-0064248.		
PR	07-NOV-1997; PR	97US-0064809.		
PR	12-NOV-1997; PR	97US-0065186.		
PR	17-NOV-1997; PR	97US-0065846.		
PR	18-NOV-1997; PR	97US-0065693.		
PR	21-NOV-1997; PR	97US-0066120.		
PR	21-NOV-1997; PR	97US-0066364.		
PR	24-NOV-1997; PR	97US-0066772.		
PR	24-NOV-1997; PR	97US-0066466.		
PR	24-NOV-1997; PR	97US-0066770.		
PR	24-NOV-1997; PR	97US-0066511.		
PR	24-NOV-1997; PR	97US-0066453.		
XX	PA (GETH) GENENTECH INC.			
XX	PI Chen J, Goddard A, Gurney AL, Pennica D, Wood WI, Yuan J;			
XX	DR WPI; 1999-229533/19.			
XX	DR N-PSDB; AAX52238.			
PT	New isolated human genes and polypeptides used in, e.g. treatment of			
PT	gastrointestinal ulceration			
XX	Claim 12; Fig 50; 320pp; English.			
PS	CC AAY13344-403 represent secreted and transmembrane human proteins.			
XX	CC The cDNA sequences are obtained from cDNA libraries, prepared from fetal lung, fetal kidney, fetal brain, fetal liver and fetal retina.			
CC	CC The encoded polypeptides have specific uses based on their homology to known polypeptides, e.g. PRO211 and PRO217 can be used for disorders associated with the preservation and maintenance of gastrointestinal			
CC	CC mucosa and the repair of acute and chronic mucosal lesions (e.g. enterocolitis, Zollinger-Ellison syndrome, gastrointestinal ulceration and congenital microvillus atrophy), skin diseases associated with abnormal keratinocyte differentiation (e.g. psoriasis, epithelial cancers such as lung squamous cell carcinoma of the vulva and gliomas), potent effects on cell growth and development, diseases related to growth or survival of nerve cells including Parkinson's disease, Alzheimer's disease, ALS, neuropathies or cancer. PRO265 can be used as for fibromodulin, e.g. for reducing dermal scarring. PRO264 can be used as a target for anti-tumor drugs. PRO533 may be used in the treatment of Usher Syndrome or Atrophia areata; PRO269 can be used as an anti-thrombotic agent; PRO287 polypeptides and portions may have therapeutic applications in wound healing and tissue repair; PRO317 can be used for treating problems of the kidney, uterus, endometrium, blood vessels, or related tissue, e.g. in the heart of genital tract.			
XX	SQ Sequence 382 AA;			
Query	Query Match 98.7%; Score 1101; DB 20; Length 382;			
	Best Local Similarity 96.3%; Pred. No. 9.5e-104;			
	Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;			
QY	1 ATGRLLS-----GOPVCRGGTQRPCYKVIFHDTSRRLNFEAKEACRDRGGQLYSI 52			
Db	22 ATGRLLSASDLRGGQPVCRGGTQRPCYKVIFHDTSRRLNFEAKEACRDRGGQLYSI 81			
QY	53 ESEDEQKLIKEKFIEENLLPSDGDWFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWYVD 112			
Db	82 ESEDEQKLIKEKFIEENLLPSDGDWFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWYVD 141			
QY	113 EPSCGSEVCVVMHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEAGE 172			
Db	142 EPSCGSEVCVVMHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAEAGE 201			
QY	173 ETELTPVLPETQEEDAKKTKEFRESREANLNAY 206			
XX	PA (GETH) GENENTECH INC.			
XX	Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;			
XX	Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;			
DR	WPI; 2001-602746/68.			
DR	N-PSDB; AAS45934.			

XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the presence of tumours, such as prostate and breast tumours, in mammals and to screen for modulators of the compounds -

XX Claim 11; Fig 20; 774pp; English.

XX Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention. The PRO polypeptides and their associated nucleic acids can be used to detect the presence of a tumour in a mammal by comparing the level of expression of a PRO polypeptide in a test sample of cells from the animal and a control sample of normal cells, whereby a higher level of expression in the test sample indicates the presence of a tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats and rabbits but are preferably human. The polypeptides can be used to stimulate tumour necrosis factor (TNF) alpha release from human blood, when contacted with it. A specific polypeptide can be used to stimulate the proliferation or differentiation of chondrocyte cells. The PRO proteins can be used to determine the presence of tumours and also susceptibility to tumour development, particularly adrenal, lung, colon, breast, prostate, rectal, cervical, or liver tumours, in mammalian subjects. The oligonucleotide probes specific for the PRO nucleic acids can be used for genetic analysis of individuals with genetic disorders.

XX Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.5e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
DB 22 ATGRLLSASDLRGGQQPYCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 112
DB 82 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 141
QY 113 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 172
DB 142 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201
QY 173 ETELTPVLPEETQEEAKTFKESREALNLAY 206
DB 202 ETELTPVLPEETQEEAKTFKESREALNLAY 235

RESULT 9
AAB80235 ID AAB80235 standard; Protein; 382 AA.
XX AC AAB80235;
XX DT 24-APR-2001 (first entry)
XX DE Human PRO234 protein.
XX Human; PRO; dermatological; antipsoriatic; cytostatic; antiinflammatory;
KW anti-parkinsonian nootropic; neuroprotective; vulnerary; cardiant;
KW anti-angiogenic; vasoconstrictive; antiasthmatic; antirheumatic; cancer;
KW anti-arthritic; anti-fertilization; antidiabetic; antiviral; diabetes;
KW ophthalmological; gene therapy; skin disease; gastrointestinal disorder;
KW ischaemia; inflammation.

XX OS sapiens.
XX PN WO200104311-A1.
XX PD 18-JAN-2001.
XX PF 22-FEB-2000; 2000WO-US04414.
XX PR 07-JUL-1999; 99US-0143048.

RESULT 10
AAB18913 ID AAB18913 standard; Protein; 273 AA.
XX AC AAB18913;

PR 26-JUL-1999; 99US-0145698.
PR 28-JUL-1999; 99US-0146222.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 05-JAN-2000; 99WO-US00219.
XX PA (GETH) GENENTECH INC.
PI Ashkenazi AJ, Botstein D, Desnoyers L, Eaton DL, Ferrara N;
PI Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME, Goddard A;
PI Godowski PJ, Grimaldi CJ, Gurney AL, Hillian RJ, Kijavlin IJ;
PI Mather JP, Pan J, Paoni NF, Roy MA, Stewart TA, Tumas D;
PI Williams PM, Wood WI;
XX WPI; 2001-081051/09.
DR N-PSDB; AAF72396.
XX PT Sixty one nucleic acids encoding PRO polypeptides which are useful in the treatment of skin diseases (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma) and neurodegenerative diseases (e.g. Alzheimer's disease) -
XX PS Claim 1; Fig 50; 393pp; English.
XX The present sequence is one of sixty one novel secreted and transmembrane PRO polypeptides. The PRO polypeptides are useful for treating skin diseases (e.g. psoriasis), cancers (e.g. lung squamous cell carcinoma), gastrointestinal disorders (e.g. enterocolitis), neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's disease), wound repair, cardiovascular disorders (e.g. ischaemia, atherosclerosis), inflammatory disorders (e.g. asthma, rheumatoid arthritis, multiple sclerosis), infertility, AIDS and diabetes and retinal disorders such as retinitis pigmentosum. The PRO nucleic acids have applications in molecular biology, including use as hybridization probes, and in chromosome and gene mapping.
SQ Sequence 382 AA;

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.5e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
DB 22 ATGRLLSASDLRGGQQPYCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 112
DB 82 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 141
QY 113 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 172
DB 142 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201
QY 173 ETELTPVLPEETQEEAKTFKESREALNLAY 206
DB 202 ETELTPVLPEETQEEAKTFKESREALNLAY 235

Query Match 98.7%; Score 1101; DB 22; Length 382;
Best Local Similarity 96.3%; Pred. No. 9.5e-104;
Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
DB 22 ATGRLLSASDLRGGQQPYCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81
QY 53 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 112
DB 82 ESEDEQKLIKEFIENLPLPSDGFDFWIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYD 141
QY 113 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 172
DB 142 EPSCGSEVCVMMHQPSAPAGIGGPFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201
QY 173 ETELTPVLPEETQEEAKTFKESREALNLAY 206
DB 202 ETELTPVLPEETQEEAKTFKESREALNLAY 235

XX DT 08-FEB-2001 (first entry)
 XX DE A novel polypeptide designated PRO1890.
 XX KW Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122; PRO1889; PRO1890; PRO1887; PRO1785; PRO4353; PRO4357; PRO4405; PRO4356; PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030; PRO4424; PRO4422; PRO4430; PRO4439; tumour; obesity; diabetes; insulinemia; kidney disorder; Bergers disease; nephropathy; Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis; Crohns disease.
 XX OS Homo sapiens.
 XX Key Location/Qualifiers
 FT Peptide 1..21 /note- "signal peptide"
 FT Modified-site 27..32 /note- "N-myristoylation site"
 FT Modified-site 66..71 /note- "N-myristoylation site"
 FT Misc-difference 76 /note- "Glu encoded by CAA"
 FT Modified-site 91..96 /note- "N-myristoylation site"
 FT Modified-site 93..98 /note- "N-myristoylation site"
 FT Modified-site 102..107 /note- "N-myristoylation site"
 FT Modified-site 109..114 /note- "N-myristoylation site"
 FT Modified-site 140..145 /note- "N-myristoylation site"
 FT Modified-site 212..217 /note- "N-myristoylation site"
 FT Domain 214..235 /note- "N-myristoylation site"
 FT Modified-site 266..269 /note- "transmembrane domain"
 FT Modified-site 266..269 /note- "cAMP- and cGMP-dependent protein kinase phosphorylation site"
 XX PN WO2000056889-A2.
 PD 28-SEP-2000.
 XX PF 01-MAR-2000; 2000WO-US05601.
 XX PR 23-MAR-1999; 99US-0125774.
 PR 23-MAR-1999; 99US-0125778.
 PR 24-MAR-1999; 99US-0125826.
 PR 31-MAR-1999; 99US-0127035.
 PR 05-APR-1999; 99US-0127706.
 PR 21-APR-1999; 99US-0130359.
 PR 27-APR-1999; 99US-0131270.
 PR 27-APR-1999; 99US-0131272.
 PR 27-APR-1999; 99US-0131291.
 PR 04-MAY-1999; 99US-0132371.
 PR 04-MAY-1999; 99US-0132379.
 PR 04-MAY-1999; 99US-0132383.
 PR 25-MAY-1999; 99US-0135750.
 PR 08-JUN-1999; 99US-0138166.
 PR 20-JUL-1999; 99US-0144791.
 PR 03-AUG-1999; 99US-0146970.
 PR 09-DEC-1999; 99US-0170262.
 XX PA (GETH) GENENTECH INC.
 XX PI Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J; Stewart TA, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2000-628263/60.
 DR N-PSDB; AAA96340.
 PR 01-DEC-1999; 99US-0170262.

XX Novel secreted and transmembrane polypeptides useful for diagnosing tumour in a mammal, for identifying agonists and antagonists of the polypeptide and for therapeutic use
 XX PS Claim 12; Fig 10; 222PP; English.
 XX CC The present sequence represents a secreted or transmembrane polypeptide.
 CC The specific sequence describes polypeptides designated PRO1484, PRO4334, PRO1122, PRO1889, PRO1890, PRO1887, PRO4353, PRO4357, PRO4405, PRO4356, PRO4352, PRO4380, PRO4354, PRO4408, PRO5737, PRO4425, PRO5990, PRO6030, PRO4424, PRO4422, PRO4430 and PRO4499. PRO1889 polypeptide is useful for diagnosing tumour in a mammal. The polypeptides, their agonists and antagonists are useful treating a condition associated with expression or activity of the polypeptide. Conditions treated include obesity, diabetes or hyper- or hypo-insulinemia. The polypeptides are capable of inducing proliferation of mammalian kidney mesangial cells and are therefore useful for treating kidney disorders associated with decreased mesangial cell function such as Bergers disease or other nephropathies associated with Schonlein-Henoch purpura, celiac disease, dermatitis herpetiformis or Crohns disease. The nucleic acids may be used to generate transgenic animals for use in development and screening of therapeutically useful reagents and also for chromosome identification
 CC and tissue typing.
 XX SQ Sequence 273 AA;
 XX Query Match 50.4%; Score 561.5; DB 21; Length 273;
 Best Local Similarity 60.1%; Pred. No. 7e-49;
 Matches 101; Conservative 25; Mismatches 35; Indels 7; Gaps 3;
 Qy 4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIESEDEQKLIK 63
 Db 23 RRVSGQKVKCFADFKHPCKMAYFHELSSRVSFQEARNLACESEGGVLLSLENEAEQKLIES 82
 Qy 64 FIEENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSIISQFRNWWYDEPSCGS 118
 Db 83 MLQNLTKPGTGISDGGDFWIGLWRNGDQT-SGACPDLYQWSDGSNSQYRNWWYDEPSCGS 141
 Qy 119 EVCVYMYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDE-KPAVP 165
 Db 142 EKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKHNICKYYPEINPTAP 189

RESULT 11
 AAU12441
 ID AAU12441 standard; Protein; 273 AA.
 XX AAU12441;
 AC AAU12441;
 XX DT 24-OCT-2001 (first entry)
 DE Human PRO1890 polypeptide sequence.
 XX KW Human secretory and transmembrane; PRO; mammalian; cancer; lung; breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha; cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle; adipocyte; A-peptide; factor VIIA; gene therapy.
 XX OS Homo sapiens.
 XX PN WO200140466-A2.
 XX PD 07-JUN-2001.
 XX PF 01-DEC-2000; 2000WO-US32678.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 09-DEC-1999; 99US-0170262.

PR	16-DEC-1999;	99WO-US30095.	Qy	119	EVCVVMMHQPSAPAGIGGPMFQWNDRCMKNNNICKYSDE-KPAVP 165
PR	20-DEC-1999;	99WO-US30911.	ID	AAB73309	standard; protein; 273 AA.
PR	20-DEC-1999;	99WO-US30999.	DB	142	ERCVVMMHQPTANPGGLGGPYLYQWNDRCMKNHNYICKYEPEINTAP 189
PR	30-DEC-1999;	99WO-US31243.			
PR	06-JAN-2000;	2000WO-US00277.			
PR	06-JAN-2000;	2000WO-US00376.	RESULT 12		
PR	11-FEB-2000;	2000WO-US03565.	AAB73309		
PR	18-FEB-2000;	2000WO-US04341.	ID	AAB73309	standard; protein; 273 AA.
PR	18-FEB-2000;	2000WO-US04342.	XX		
PR	22-FEB-2000;	2000WO-US04414.	AC	AAB73309;	
PR	24-FEB-2000;	2000WO-US04914.	XX	22-MAY-2001	(first entry)
PR	24-FEB-2000;	2000WO-US05004.	DT		
PR	01-MAR-2000;	2000WO-US05601.	XX		
PR	20-MAR-2000;	2000WO-US07377.	DE		
PR	21-MAR-2000;	2000WO-US07532.	XX		
PR	30-MAR-2000;	2000WO-US08439.	KW		
PR	17-MAY-2000;	2000WO-US13705.	KW		
PR	22-MAY-2000;	2000WO-US14042.	KW		
PR	30-MAY-2000;	2000WO-US14941.	XX		
PR	02-JUN-2000;	2000WO-US15264.	OS		
PR	10-NOV-2000;	2000WO-US30873.	XX		
PA	(GETH) GENENTECH INC.		PN	WO200112811-A1.	
XX			XX		
PI	Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;		PD	22-FEB-2001.	
PI	Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;		XX		
PI	Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;		PF	11-AUG-2000; 2000WO-US22065.	
XX			XX		
DR	WPI; 2001-408281/43.		PR	12-AUG-1999; 99US-0148935.	
DR	N-PSDB; AAS21513.		XX		
XX			PA	(UROG -) UROGENEWS INC.	
PT	Isolated, secretory and transmembrane PRO polypeptide used to detect other PRO polypeptides, link bioactive molecules to cells expressing PRO polypeptides, and detect the presence of mammalian tumours e.g. lung, breast, prostate, cervical -		PI	Afar DEH, Hubert RS, Jakobovits A, Raitano AB;	
PT			XX	WPI; 2001-211222/21.	
PT			XX	DR N-PSDB; AAF76004.	
PS	Claim 12; Fig 540; 813PP; English.		XX	PT	New PC-LECTIN polynucleotide encoding a transmembrane antigen over expressed in human prostate cancer, useful for the prognosis, diagnosis and treatment of prostate cancer
XX			XX	PT	and treatment of prostate cancer
CC	AAU12172-AAU12446 represent novel human secretory and transmembrane PRO polypeptides. The PRO polypeptides are useful to detect other PRO polypeptides, to link bioactive molecules to cells expressing PRO polypeptides, to modulate biological activities of cells expressing PRO polypeptides, and to detect the presence of mammalian lung, colon, breast, prostate, rectal, cervical or liver tumours by comparing PRO polypeptide expression in a cell sample to that in a control sample. Some of the 275 sequences are also useful to stimulate the release of tumour necrosis factor-alpha (TNF-alpha) from human blood, the proliferation or differentiation of chondrocytes, the proliferation or gene expression in pericyte cells, the release of proteoglycans from cartilage, the proliferation of inner ear utricular supporting cells or of T-lymphocytes, the release of a cytokine from peripheral blood monocytes (PBMCs), or the proliferation of endothelial cells. Some of the PRO polypeptides may modulate glucose or free fatty acid uptake by skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide to factor VIIA. The PRO polypeptides can be used in assays to identify molecules involved in binding interactions. The polynucleotides encoding PRO polypeptides can be used to generate probes, antisense RNA/DNA, transgenic or knock out animals and can be used in gene therapy.		PS	Claim 1; Fig 1A-D; 116PP; English.	
CC			XX	CC	The invention relates to a novel human C-type lectin transmembrane antigen, PC-LECTIN (AAB73309) and cDNA encoding it (AAF76004). The expression of the human PC-LECTIN gene is normally restricted to the testis, but is highly overexpressed in prostate cancer. PC-LECTIN expression is higher in androgen-independent prostate tumours compared with androgen-independent prostate tumours, and expression is therefore likely to be dependent on the presence of androgen. Human PC-LECTIN therefore represents a diagnostic and therapeutic target for prostate cancer, particularly androgen-dependent prostate cancer.
CC			CC	CC	Human PC-LECTIN exhibits homology to hamster layilin (44.9% identity over a 265 residue overlap), but is not thought to be the human orthologue of layilin, as diverges significantly in a key functional domain proposed for the layilin protein. Human PC-LECTIN or an immunogenic portion thereof, a vector encoding PC-LECTIN, a PC-LECTIN antisense nucleotide, a PC-LECTIN nucleotide-targetted ribozyme, or an anti-PC-LECTIN antibody may be used to prepare a composition for treating a patient with a cancer, particularly prostate cancer, but also breast, bladder, lung, bone, colon, pancreatic, testicular, cervical or ovarian cancers that express PC-LECTIN. PC-LECTIN proteins are also useful for diagnosing the presence of cancer. PC-LECTIN antibodies and nucleotides are useful in the treatment (e.g., antisense therapy), diagnosis and/or prognosis of prostate cancer and other PC-LECTIN-expressing cancers. PC-LECTIN antibodies may also be used as drug targeting agents. The PC-LECTIN nucleotides and proteins may additionally be used in drug discovery to identify molecules that modulate PC-LECTIN function or expression. The present sequence represents human PC-LECTIN.
CC			CC	CC	XX
CC			CC	CC	Sequence 273 AA;
CC			CC	CC	Query Match 50.4%; Score 561.5; DB 22; Length 273;
CC			CC	CC	Best Local Similarity 60.1%; Pred. No. 7e-49; Mismatches 25; Indels 7; Gaps 3;
CC			CC	CC	Matches 101; Conservative 25; Mismatches 35; Indels 7; Gaps 3;
Qy	4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSISEDEQKLIEK 63		Qy	4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSISEDEQKLIEK 63	
Db	23 RVVSGQVKCFADFRHPCYKMAVFHELSSRVSFQEALACESEGGVLLSLENEAEQKLIES 82		Db	23 RVVSGQVKCFADFRHPCYKMAVFHELSSRVSFQEALACESEGGVLLSLENEAEQKLIES 82	
Qy	64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGGSISQFRNWYDEPSCGS 118		Qy	64 FIENLLP-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGGSISQFRNWYDEPSCGS 118	
Db	83 MLQNLTKPQGTGISDGDGEWIGLWRNGDQT-SGACPDLYQWSDGSNSQYRNWYDEPSCGS 141		Db	83 MLQNLTKPQGTGISDGDGEWIGLWRNGDQT-SGACPDLYQWSDGSNSQYRNWYDEPSCGS 141	
			SQ	Sequence 273 AA;	
			Query Match	50.4%; Score 561.5; DB 22; Length 273;	

Best Local Similarity 60.1%; Pred. No. 7e-49; Matches 101; Conservative		SQ Sequence 273 AA;	
Qy 4 RLLSGQPVCRGGTQRPCYKVIVFHDTSRRLNFEAKEACRRDGQLVSISEDEQKLIEK 63		Query Match 50.4%; Best Local Similarity' 60.1%; Pred. No. 7e-49; Matches 101; Conservative	Length 273;
Db 23 RVVSGQKVCFADFKHPCYKMAFYHELSSRVSFQEARNLACESEGVLSSLENEAEQKLIES 82		25; Mismatches 35; Indels 7; Gaps 3;	
Qy 64 FIENLLP----SDGDFWIGLRRREEKOSNINSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118		Qy 4 RLLSGQPVCRGGTQRPCYKVIVFHDTSRRLNFEAKEACRRDGQLVSISEDEQKLIEK 63	
Db 83 MLQNLTKPQGTGISDGDFWIGLWRNGDQT-SGACPDLQWSDGSNSQYRNWYTDPEPSCGS 141		Db 23 RVVSGQKVCFADFKHPCYKMAFYHELSSRVSFQEARNLACESEGVLSSLENEAEQKLIES 82	
Qy 119 EVCVVMMYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDE-KPAVP 165		Qy 64 FIENLLP----SDGDFWIGLRRREEKOSNINSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118	
Db 142 EKCVVMMYHQPTANPGLGGPYLYQWNDDRCNMKHNICKYEEPEINPTAP 189		Db 83 MLQNLTKPQGTGISDGDFWIGLWRNGDQT-SGACPDLQWSDGSNSQYRNWYTDPEPSCGS 141	
RESULT 13 AAB87609 standard; protein; 273 AA.		RESULT 14 ABB95554 standard; protein; 273 AA.	
XX AAB87609;	XX ABB95554;	XX ABB95554;	XX ABB95554;
AC DT 15-MAY-2001 (first entry)	AC DT 19-JUL-2002 (first entry)	AC DT 19-JUL-2002 (first entry)	AC DT 19-JUL-2002 (first entry)
XX Human PRO1890.	XX Human angiogenesis related protein PRO1890 SEQ ID NO: 264.	XX Human angiogenesis related protein PRO1890 SEQ ID NO: 264.	XX Human angiogenesis related protein PRO1890 SEQ ID NO: 264.
KW DE	KW DE	KW DE	KW DE
XX Homo sapiens.	XX atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder; cardiant; cytostatic; antiangiogenic; hypotensive; pulmonary; antiarteriosclerotic.	XX atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder; cardiant; cytostatic; antiangiogenic; hypotensive; pulmonary; antiarteriosclerotic.	XX atherosclerosis; cardiac hypertrophy; gene therapy; endothelial disorder; cardiant; cytostatic; antiangiogenic; hypotensive; pulmonary; antiarteriosclerotic.
OS XX	OS XX	OS XX	OS XX
PN WO200116318-A2.	PN WO200208284-A2.	PN WO200208284-A2.	PN WO200208284-A2.
PD XX	PD XX	PD XX	PD XX
XX 08-MAR-2001.	XX 31-JAN-2002.	XX 31-JAN-2002.	XX 31-JAN-2002.
PF XX	PF XX	PF XX	PF XX
XX 24-AUG-2000; 2000WO-US233328.	XX 09-JUL-2001; 2001WO-US21735.	XX 09-JUL-2001; 2001WO-US21735.	XX 09-JUL-2001; 2001WO-US21735.
XX PR 01-SEP-1999; 99WO-US20111.	XX PR 20-JUL-2000; 2000US-219556P.	XX PR 20-JUL-2000; 2000US-219556P.	XX PR 20-JUL-2000; 2000US-219556P.
PR 15-SEP-1999; 99WO-US21090.	PR 25-JUL-2000; 2000US-220624P.	PR 25-JUL-2000; 2000US-220624P.	PR 25-JUL-2000; 2000US-220624P.
PR 07-DEC-1999; 99US-0169495.	PR 25-JUL-2000; 2000US-220664P.	PR 25-JUL-2000; 2000US-220664P.	PR 25-JUL-2000; 2000US-220664P.
PR 09-DEC-1999; 99US-0170262.	PR 28-JUL-2000; 2000WO-US20710.	PR 28-JUL-2000; 2000WO-US20710.	PR 28-JUL-2000; 2000WO-US20710.
PR 11-JAN-2000; 2000US-0175481.	PR 02-AUG-2000; 2000US-222695P.	PR 02-AUG-2000; 2000US-222695P.	PR 02-AUG-2000; 2000US-222695P.
PR 18-FEB-2000; 2000WO-US04341.	PR 17-AUG-2000; 2000US-0643657.	PR 17-AUG-2000; 2000US-0643657.	PR 17-AUG-2000; 2000US-0643657.
PR 18-FEB-2000; 2000WO-US04342.	PR 23-AUG-2000; 2000WO-US23522.	PR 23-AUG-2000; 2000WO-US23522.	PR 23-AUG-2000; 2000WO-US23522.
PR 22-FEB-2000; 2000WO-US04414.	PR 24-AUG-2000; 2000WO-US23328.	PR 24-AUG-2000; 2000WO-US23328.	PR 24-AUG-2000; 2000WO-US23328.
PR 01-MAR-2000; 2000WO-US05601.	PR 07-SEP-2000; 2000US-230978P.	PR 07-SEP-2000; 2000US-230978P.	PR 07-SEP-2000; 2000US-230978P.
PR 03-MAR-2000; 2000US-0187202.	PR 15-SEP-2000; 2000US-000000P.	PR 15-SEP-2000; 2000US-000000P.	PR 15-SEP-2000; 2000US-000000P.
PR 25-APR-2000; 2000US-0199397.	PR 18-SEP-2000; 2000US-0664610.	PR 18-SEP-2000; 2000US-0664610.	PR 18-SEP-2000; 2000US-0664610.
PR 22-MAY-2000; 2000WO-US14042.	PR 10-NOV-2000; 2000US-0747259.	PR 10-NOV-2000; 2000US-0747259.	PR 10-NOV-2000; 2000US-0747259.
PR 05-JUN-2000; 2000US-0209832.	PR 20-DEC-2000; 2000WO-US34956.	PR 20-DEC-2000; 2000WO-US34956.	PR 20-DEC-2000; 2000WO-US34956.
XX PA (GETH) GENENTECH INC.	PR 22-JAN-2001; 2001US-0767609.	PR 22-JAN-2001; 2001US-0767609.	PR 22-JAN-2001; 2001US-0767609.
XX PR 28-FEB-2001; 2001US-0796498.	PR 28-FEB-2001; 2001US-0796498.	PR 28-FEB-2001; 2001US-0796498.	PR 28-FEB-2001; 2001US-0796498.
PI 01-MAR-2001; 2001WO-US06666.	PR 01-MAR-2001; 2001WO-US06666.	PR 01-MAR-2001; 2001WO-US06666.	PR 01-MAR-2001; 2001WO-US06666.
PI 09-MAR-2001; 2001US-0802706.	PR 14-MAR-2001; 2001US-0802706.	PR 14-MAR-2001; 2001US-0802706.	PR 14-MAR-2001; 2001US-0802706.
PI 14-MAR-2001; 2001US-0808689.	PR 22-MAR-2001; 2001US-0816744.	PR 22-MAR-2001; 2001US-0816744.	PR 22-MAR-2001; 2001US-0816744.
PI 22-MAR-2001; 2001US-0816744.	PR 05-APR-2001; 2001US-0828366.	PR 05-APR-2001; 2001US-0828366.	PR 05-APR-2001; 2001US-0828366.
XX PR 10-MAY-2001; 2001US-0854208.	PR 10-MAY-2001; 2001US-0854208.	PR 10-MAY-2001; 2001US-0854208.	PR 10-MAY-2001; 2001US-0854208.
PS PR 10-MAY-2001; 2001US-0866028.	PR 25-MAY-2001; 2001US-0866028.	PR 25-MAY-2001; 2001US-0866028.	PR 25-MAY-2001; 2001US-0866028.
XX PR 25-MAY-2001; 2001US-0866034.	PR 25-MAY-2001; 2001US-0866034.	PR 25-MAY-2001; 2001US-0866034.	PR 25-MAY-2001; 2001US-0866034.
The present sequence is a human PRO polypeptide (secreted and transmembrane). The PRO protein, and PRO agonists or anti-PRO antibodies are useful for preparation of a medicament useful in the treatment of a condition which is responsive to the PRO protein, agonists, antagonists or anti-PRO antibodies. The PRO protein may also be employed as molecular weight markers for protein electrophoresis. The PRO coding sequence has applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping.			

PR	25-MAY-2001; 2001WO-US17092.	DE	Human PRO1890 protein sequence SEQ ID NO:264.
PR	30-MAY-2001; 2001US-087054.	XX	Human; angiogenesis; cardiant; cytostatic; antiangiogenic; hypotensive;
PR	30-MAY-2001; 2001WO-US17443.	KW	vulnerary; antiarteriosclerotic; PRO antagonist; PRO antagonist; trauma;
PR	01-JUN-2001; 2001WO-US17800.	KW	gene therapy; cardiovascular disorder; endothelial disorder; cancer;
PR	20-JUN-2001; 2001WO-US19692.	KW	angiogenic disorder; cardiac hypertrophy; atherosclerosis; hypertension;
PR	28-JUN-2001; 2001WO-US00000.	KW	age-related macular degeneration; arterial restenosis; angina;
XX		KW	rheumatoid arthritis; myocardial infarction; thrombophlebitis;
PA	(GETH) GENENTECH INC.	KW	lymphangitis; tumour angiogenesis; breast carcinoma; liver carcinoma;
PA	(BAKE/) BAKER K P.	KW	wound healing; chromosome mapping; gene mapping.
PA	(FERR/) FERRARA N.	XX	
PA	(GERB/) GERBER H.	XX	
PA	(GERR/) GERRITSEN M E.	XX	
PA	(GODD/) GODDARD A.	OS	
PA	(GODO/) GODOWSKI P J.	PN	
PA	(GURN/) GURNEY A L.	XX	
PA	(HILL/) HILLIAN K J.	XX	
PA	(MARS/) MARSTERS S A.	PD	
PA	(PANJ/) PAN J.	XX	
PA	(PAON/) PAONI N F.	PF	
PA	(STEP/) STEPHAN J F.	XX	
PA	(WATA/) WATANABE C K.	PR	23-JUN-2000; 2000US-213637P.
PA	(WILL/) WILLIAMS P M.	PR	20-JUL-2000; 2000US-219556P.
PA	(WOOD/) WOOD W I.	PR	25-JUL-2000; 2000US-220624P.
XX		PR	25-JUL-2000; 2000US-220664P.
PI	Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Hillian KJ, Marsters SA, Pan J, Paoni NF;	PR	28-JUL-2000; 2000WO-US20710.
PI	Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;	PR	02-AUG-2000; 2000US-222695P.
PI	XX	PR	17-AUG-2000; 2000US-0643657.
DR	WPI; 2002-171999/22.	PR	23-AUG-2000; 2000WO-US23522.
DR	N-PSDB; ABL95692.	PR	24-AUG-2000; 2000WO-US23328.
XX		PR	07-SEP-2000; 2000US-230978P.
PT	One hundred and eighty seven nucleic acids encoding PRO polypeptides, useful in diagnosis and treatment of cardiovascular (e.g. myocardial infarction), endothelial or angiogenic disorders in a mammal -.	PR	18-SEP-2000; 2000US-0664610.
PT	XX	PR	24-OCT-2000; 2000US-242922P.
PT	XX	PR	08-NOV-2000; 2000US-0709238.
PS	XX	PR	08-NOW-2000; 2000WO-US30952.
PS	XX	PR	10-NOV-2000; 2000WO-US30873.
XX		PR	01-DEC-2000; 2000US-0665350.
CC		PR	20-DEC-2000; 2000US-0747259.
CC		PR	2000WO-US34956.
CC		PR	22-JAN-2001; 2001US-0767609.
CC		PR	28-FEB-2001; 2001US-0796498.
CC		PR	28-FEB-2001; 2001WO-US06520.
CC		PR	01-MAR-2001; 2001WO-US06666.
CC		PR	09-MAR-2001; 2001US-0802706.
CC		PR	14-MAR-2001; 2001US-0808689.
CC		PR	22-MAR-2001; 2001US-0816744.
CC		PR	05-APR-2001; 2001US-0828366.
CC		PR	10-MAY-2001; 2001US-0854208.
CC		PR	10-MAY-2001; 2001US-0854280.
CC		PR	25-MAY-2001; 2001US-0866028.
CC		PR	25-MAY-2001; 2001US-0866034.
Qy	4 RLLSGQPVCRGGTQRPCYKVIVFHDTSRRLNFEAKEACRRDGGOLVSISEDEQKLIEK	3	PR 25-MAY-2001; 2001WO-US17092.
Db	23 RVVSGQKVCADFHKPCYKMAVFHELSSRVSFQEALACESEGVLSSLENEAEQKLIES	82	PR 30-MAY-2001; 2001US-0870574.
Qy	64 FIENLLP-----SDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS	118	PR 01-JUN-2001; 2001WO-US17443.
Db	83 MLQNLTRPKGTGTDGDFWIGLWRNGDQT-SGACPDLYQWSDGSNSQYRNWYTDPEPSCGS	141	PR 01-JUN-2001; 2001WO-US17800.
Qy	119 EVCVVMYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDE-KPAVP 165	XX	PA (GETH) GENENTECH INC.
Db	142 EKCVVVMYHQPTANPGLGGPYLQWNDDRCNMKHNICKYEPINEINTAP 189	XX	Baker KP, Ferrara N, Gerber H, Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Hillian KJ, Marsters SA, Pan J, Paoni NF;
XX		XX	Stephan JF, Watanabe CK, Williams PM, Wood WI, Ye W;
PS		XX	WPI; 2002-090516/12.
AC	ABB84948	XX	DR N-PSDB; ABL88203.
AC	ABB84948 standard; Protein; 273 AA.	XX	XX
DT	16-MAY-2002 (first entry)	XX	One hundred and eighty seven nucleic acids encoding PRO polypeptides, useful in diagnosis and treatment of cardiovascular (e.g. myocardial infarction), endothelial or angiogenic disorders in a mammal -.
PS	Claim 11; Fig 264; 565pp; English.	XX	PT PT
XX	ABL88072 to ABL88258 encode the PRO proteins given in ABB84817 to ABB8503. The PRO proteins and polynucleotides have cardiant, cytostatic, antiangiogenic, hypotensive, vulnerary and antiarteriosclerotic	CC CC	CC

activities, and can be used in gene therapy. The PRO polynucleotides, proteins, agonists and antagonists are useful for treating or diagnosing a cardiovascular, endothelial or angiogenic disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-related macular degeneration, atherosclerosis, hypertension, arterial restenosis, rheumatoid arthritis, angina, myocardial infarctions, thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast carcinoma and liver carcinoma) and wound healing. The PRO polynucleotides have applications in molecular biology, including use as hybridisation probes, and in chromosome and gene mapping. ABL88259 to ABL88267 represent primers and probes used in the exemplification of the present invention.

Sequence 273 AA;

Query Match 50.4%; Score 561.5; DB 23; Length 273;
Best Local Similarity 60.1%; Pred. No. 7e-49;
Matches 101; Conservative 25; Mismatches 35; Indels 7; Gaps 3;

Qy 4 RLLSGQPVCRGGTQRPCYKVIYFHDTSSRRLNFEAKEACRRDGQLVSVIESEDEQKLEK 63
Db 23 RVVSGQKVCFADFRHPCYKMAVFHELSSRVSFQEARNLACESEGVLSSLENEAEQKLIES 82

Qy 64 FIEPLL-----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSIQFRNWYVDEPSCGS 118
Db 83 MLQNLTKPGTGISDGDFWIGLWRNGDGQT-SGACPDLQWSDGSNSQYRNWYVDEPSCGS 141

Qy 119 EVCVVMYHOPSAAPAGIGGPMQWNDDRCNMKNNFICKYSDE-KPAVP 165
Db 142 EKCVVMYHOPTANPGLGGPYLYQWNDDRCNMKHNICKYEPEinPTAP 189

Search completed: May 20, 2003, 17:05:07
Job time : 77 secs

GenCore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:03:50 ; search time 15 Seconds
 (without alignments)
 404.075 Million cell updates/sec

Title: US-09-887-855-5
 Perfect score: 1115
 Sequence: 1 ATGRLLSGQPVCRGGTQRPC. EEDAKKTFKESREAALNLAY 206

Scoring table: BLOSUM62
 Gapext 0.5
 Gapop 10.0

Sequence 8, App11
 Sequence 11, App1
 Sequence 14, App1
 Sequence 17, App1
 Sequence 20, App1
 Sequence 2, App11
 Sequence 5, App11
 Patent No. 5514582
 Sequence 13, App1
 Sequence 9, App11
 Sequence 9, App11
 Sequence 6, App11
 Sequence 16, App1
 Sequence 37, App1
 Sequence 2, App11
 Patent No. 5514582
 Sequence 2, App11
 Sequence 2, App11

Total number of hits satisfying chosen parameters: 362574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Maximum Match 100%
Listing first 45 summaries

```
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*
```

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query	Match	Length	DB ID	Description
1	178.5	16.0	1455	3	US-08-840-062-5	Sequence 5, Appl 11
2	176.5	15.8	1479	3	US-08-840-062-4	Sequence 4, Appl 11
3	174	15.6	1257	1	US-08-340-428B-49	Sequence 49, Appl 11
4	174	15.6	2409	6	5180808-2	Patent No. 5180808
5	170.5	15.3	1479	3	US-08-840-062-2	Sequence 2, Appl 11
6	152	13.6	912	5	PCT-US95-03747-2	Sequence 2, Appl 11
7	151	13.5	197	4	US-09-602-877A-99	Sequence 99, Appl 11
8	149	13.4	174	1	US-07-641-971B-1	Sequence 1, Appl 11
9	149	13.4	174	1	US-07-781-248A-1	Sequence 1, Appl 11
10	149	13.4	320	1	US-08-365-103B-10	Sequence 10, Appl 11
11	149	13.4	321	1	US-08-365-103B-8	Sequence 8, Appl 11
12	148.5	13.3	652	2	US-08-751-305-2	Sequence 2, Appl 11
13	144.5	13.0	404	4	US-09-517-605-2	Sequence 2, Appl 11
14	143.5	12.9	372	2	US-08-513-278-4	Sequence 4, Appl 11
15	143.5	12.9	372	6	5514582-4	Patent No. 5514582
16	142	12.7	1487	3	US-08-840-062-7	Sequence 7, Appl 11
17	141.5	12.7	125	3	US-08-722-126A-7	Sequence 7, Appl 11
18	141.5	12.7	125	5	PCT-US95-04258-7	Sequence 7, Appl 11
19	141.5	12.7	287	1	US-08-365-103B-4	Sequence 4, Appl 11
20	141.5	12.7	300	1	US-08-365-103B-6	Sequence 6, Appl 11
21	141.5	12.7	327	1	US-08-365-103B-2	Sequence 2, Appl 11
22	138.5	12.4	117	6	5514582-7	Patent No. 5514582
23	138.5	12.4	119	1	US-08-340-539A-12	Sequence 12, Appl 11
24	137.5	12.3	110	6	5514582-12	Patent No. 5514582
25	137.5	12.3	123	6	5514582-19	Sequence 39, Appl 11
26	137.5	12.3	492	4	US-09-724-864-39	Patent No. 5514582
27	135	12.1	238	4	US-09-111-470-8	Sequence 8, Appl 11

28	134.5	12.1	128	4	US-09-535-521-8	Sequence 8, Appl 1
29	134.5	12.1	139	4	US-09-535-521-11	Sequence 11, Appl 1
30	134.5	12.1	141	4	US-09-535-521-14	Sequence 14, Appl 1
31	134.5	12.1	187	4	US-09-535-521-17	Sequence 17, Appl 1
32	134.5	12.1	208	4	US-09-535-521-20	Sequence 20, Appl 1
33	134.5	12.1	292	4	US-09-535-521-2	Sequence 2, Appl 1
34	134.5	12.1	292	4	US-09-535-521-5	Sequence 5, Appl 1
35	133.5	12.0	110	6	5514582-9	Patent No. 5514582
36	133.5	12.0	119	1	US-08-340-539A-13	Sequence 13, Appl 1
37	133.5	12.0	133	1	US-07-893-929A-9	Sequence 9, Appl 1
38	133.5	12.0	133	5	PCT-US92-10344-9	Sequence 9, Appl 1
39	133.5	12.0	287	4	US-09-111-470-6	Sequence 6, Appl 1
40	132.5	11.9	119	1	US-08-340-539A-16	Sequence 16, Appl 1
41	132.5	11.9	120	1	US-08-274-661B-37	Sequence 37, Appl 1
42	132.5	11.9	372	2	US-08-513-278-2	Sequence 2, Appl 1
43	132.5	11.9	372	6	5514582-2	Patent No. 5514582
44	132.5	11.9	385	1	US-08-340-539A-2	Sequence 2, Appl 1
45	132.5	11.9	385	2	US-08-461-592B-2	Sequence 2, Appl 1

ALIMENTC

RESULT 1
US-08-84
; Sequence
; Patent
; GENERAL
; APP APP
; APP APP
; TIT TIT
; NUM NUM
; COR COR
; A A
; S S
; C C
; C C
; Z Z
; COM COM
; M M
; C C
; O O
; S S
; CUR CUR
; A A
; F F
; C C
; ATT ATT
; N N
; R R
; R R
; TEL TEL
; T T
; T T
; T T
; INFOR INFOR
; SEQ SEQ
; L L
; T T
; T T
; T T
; US-08-84
; US-08-84
Query 2
Best L
Matche
; QY QY
; DDb 80
; QY QY 8

Digitized by srujanika@gmail.com

Db 865 ISMDKK-----FIWMDGSKVDFVAWATGEPNFANDDENCVNY---TNSGF---- 907
 Qy 139 MFQWNDRCNMKNNFICK---YSDEKPAVPSREAEGETELTPVLPEETQE----- 187
 Db 908 ---WNDINCGYPNPFICQRHNSSTINATAMP-----TPTTPGGCKEGWHLYKNK 953
 Qy 188 -----EDAKKTFKESREAAALNL 204
 Db 954 CFKIFGFANEEKKSWQDARACKGL 978

RESULT 2
 US-08-840-062-4
 ; Sequence 4, Application US/08840062
 ; Patent No. 611977
 ; GENERAL INFORMATION:
 ; APPLICANT: LASKY, LAURENCE A.
 ; TITLE OF INVENTION: TYPE C LECTINS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/840,062
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Drieger, Ginger R.
 ; REGISTRATION NUMBER: 33,055
 ; REFERENCE/DOCKET NUMBER: P1019R1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-3216
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1479 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ; US-08-840-062-4

Query Match 15.8%; Score 176.5; DB 3; Length 1479;
 Best Local Similarity 30.4%; Pred. No. 1.4e-09; Mismatches 25; Indels 35; Caps 7;

Qy 7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAEKACRRDGQLVSISEDEQKLIKEKFILE 66
 Db 385 SWQPF----QGHCYRL----QAEKRSWQESKKAQLRGGGDLVSIHSMMAELEFITKQIK 434

Qy 67 NLLPSDGFDWIGLRRREEQNSNSTACQDLIYAWTDGSISQFRNHYVDEPS---CGSEVCVV 123
 Db 435 QEVE--ELWIGL----NDLKLQMNFEWSDGSLSVFSFTHWHPFEPNNFRDSLEDCT 483

Qy 124 MYHQPSAPAGIGGPFQWQNDRCNMKNNFICKYSDEKPAVPSREAEQ 171
 Db 484 IW----GPEG----RWNDSPCNQSLPSICKKAGQLSQGAAEDHG 520

RESULT 3
 US-08-340-428B-49
 ; Sequence 49, Application US/08340428B
 ; Patent No. 5648465
 ; GENERAL INFORMATION:

; APPLICANT: MARGOLIS, Richard U.
 ; APPLICANT: FAUCH, Uwe
 ; APPLICANT: MARGOLIS, Renee K.
 ; TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
 ; NEUROCAN AS A CHONDROITIN SULFATE PROTEOGLYCAN
 ; NUMBER OF SEQUENCES: 49
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Browdy and Neimark
 ; STREET: 419 Seventh Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: U.S.A.
 ; ZIP: 20004
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/340,428B
 ; FILING DATE: 14 NO. 5648465ember 1994
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/922,911
 ; FILING DATE: 03 August 1992
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Browdy, Roger L.
 ; REGISTRATION NUMBER: 25,618
 ; REFERENCE/DOCKET NUMBER: Margolis-1A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-628-5197
 ; TELEFAX: 202-737-3528
 ; INFORMATION FOR SEQ ID NO: 49:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1257 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-340-428B-49

Query Match 15.6%; Score 174; DB 1; Length 1257;
 Best Local Similarity 30.8%; Pred. No. 2e-09; Mismatches 49; Indels 32; Caps 6;

Qy 17 QRPCYKVIYFHDTSRRLNFEAEKACRRDGQLVSISEDEQKLIKEKFILE 76
 Db 1037 QGHCYR--YF--AHRRAWDAERDCCRRAGLTSVHSPEEKFINSF----GHENSW 1085

Qy 77 IGLRRREEQNSNSTACQDLIYAWTDGSISQFRNHYVDEPS---CGSEVCVVYHQSAPAG 133
 Db 1086 IGLNDRTVERD----FQWTDNTGLQYENREKQPDNFFAGGEDCYYMVVAHENG--- 1134

Qy 134 IGGPYMFQWQNDRCNMKNNFICK 156
 Db 1135 -----RWNDVPCNYNLPYVCK 1150

RESULT 4
 5180808-2
 ; Patent No. 5180808
 ; APPLICANT: RUOSLAHTI, ERKKI I.
 ; TITLE OF INVENTION: VERSICAN CORE PROTEIN, NUCLEIC ACID
 ; SEQUENCES ENCODING THE SAME, NUCLEIC ACID PROBES, ANTI-VERSICAN
 ; ANTIBODIES, AND METHODS OF DETECTING THE SAME
 ; NUMBER OF SEQUENCES: 4
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/441,179
 ; FILING DATE: 27-NOV-1989
 ; SEQ ID NO:2:
 ; LENGTH: 2409
 ; 5180808-2

Query Match 15.6%; Score 174; DB 6; Length 2409;
 Best Local Similarity 28.5%; Pred. No. 4.9e-09;
 Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

Qy 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKEFIENLLPSDGF- 75
 Db 2190 QGQCYK--YF---AHRRTWDAARECRLQGAHLTSILSHEEQMFVNVR----GHDYQ 2237

Qy 76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWFDEP---SCGSEVCVVMYHQPSAP 131
 Db 2238 WIGL--NDKMFHDFRWTDGSTLQYENWRPNQPSFFSAGEDCVVIIWHENG-- 2287

Qy 132 AGIGGPMFQWNDRCNMKNNFICKYS---DEKPAVPSREAGE 172
 Db 2288 -----QWNDDVPCNYHLTYTCKKGTVACGQPPVVENAKTTFGK 2323

RESULT 5
 US-08-840-062-2
 ; Sequence 2, Application US/08840062
 ; Patent No. 6117977
 ; GENERAL INFORMATION:
 ; APPLICANT: LASKY, LAURENCE A.
 ; TITLE OF INVENTION: TYPE C LECTINS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/840,062
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dreger, Ginger R.
 ; REGISTRATION NUMBER: 33,055
 ; REFERENCE/DOCKET NUMBER: P1019R1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-3216
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1479 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ; US-08-840-062-2

Query Match 15.3%; Score 170.5; DB 3; Length 1479;
 Best Local Similarity 31.4%; Pred. No. 5.9e-09;
 Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

Qy 7 SGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKEFIE 66
 Db 384 SWQPF----QGHCYRL---QAEKRSQESKRACLRGGDLLSIHSMMAELEFITKQIK 433

Qy 67 NLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWFDEP---CGSEVCVV 123
 Db 434 QEVE--ELWIGL----NDLKLQMNFNEDSGSLVSFTWHWPFPENNFRDSLEDCVY 482

Qy 124 MYHQPSAPAGIGGPMFQWNDDRCNMKNNFICK 156
 Db 483 IW----GPEG----RNDSPCNQSLSTICK 504

RESULT 6
 PCT-US95-03747-2
 ; Sequence 2, Application PC/TUS9503747
 ; GENERAL INFORMATION:
 ; APPLICANT: LA JOLLA CANCER RESEARCH FOUNDATION
 ; TITLE OF INVENTION: Brevican, A Glial Cell Proteoglycan
 ; NUMBER OF SEQUENCES: 3
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Campbell and Flores
 ; STREET: 4370 La Jolla Village Drive, Suite 700
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92122
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/03747
 ; FILING DATE: 27-MAR-1995
 ; CLASSIFICATION:
 ; NAME: Imbra, Richard J.
 ; REGISTRATION NUMBER: 37,643
 ; REFERENCE/DOCKET NUMBER: FP-LJ 1453
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 535-9001
 ; TELEFAX: (619) 535-8949
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 912 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US95-03747-2

Query Match 13.6%; Score 152; DB 5; Length 912;
 Best Local Similarity 28.8%; Pred. No. 2.6e-07;
 Matches 47; Conservative 17; Mismatches 61; Indels 38; Gaps 8;

Qy 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIKEFIEENLLPSDGF 76
 Db 697 QGACYK----HFSARR-SWEEAENKCRMGAHLASISTPEEQOLFINNRYREYQ----W 745

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWFDEP---CGSEVCVV---YHQPSAPA 132
 Db 746 IGL----NDRTIEGDFLWSDGVPULLYENWNPGQPSYFLSGENCVVVMWHDQG-- 794

Qy 133 GIGGPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAE 171
 Db 795 -----QWSDVPCNTHLSYTKMGLVSGGPPPEFLAEVFG 829

RESULT 7
 US-09-602-877A-99
 ; Sequence 99, Application US/09602877A
 ; Patent No. 6432707
 ; GENERAL INFORMATION:
 ; APPLICANT: Reed, Steven G.
 ; ATTORNEY: Xu, Jiangchun
 ; ATTORNEY: Dillon, Davin C.
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER
 ; FILE REFERENCE: 210121.446C5
 ; CURRENT APPLICATION NUMBER: US/09/602,877A
 ; CURRENT FILING DATE: 2000-06-22
 ; NUMBER OF SEQ ID NOS: 107
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0
 ; SEQ ID NO 99

LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapien
us-09-602-877A-99

Query Match 13.5%; Score 151; DB 4; Length 197;
Best Local Similarity 25.0%; Pred. No. 4.2e-08;
Matches 38; Conservative 26; Mismatches 64; Indels 24; Gaps 4;

Qy 9 QPVCRGGTQ--RPCYKVIYFHDTSSRLNFEAKEACRRDGGOLVSIESTEDEQKLIKEKFIE 66
Db 65 QTVCLRGTKVHKKCYLA---SEGLKHFHEANEDCISKGGLIVPRNSDEINALQDYK 119

Qy 67 NLLPSDGFDFWIGLARRKEEKQNSNSTACQDLYAWTDGSISQFRNWWYDEPSCGSEYCVVMMYH 126
Db 120 RSLPGVNDFWLGI-----NDMVTECKFDVNGIAISFLNWDRAQPNNGKRENCVLFS 171

Qy 127 QPSAPAGIGGGPYMFQWNDRCNMKNNFICKYS 158
Db 172 QSA-----QGKWSDEACRSSKRYICEFT 194

RESULT 8
US-07-641-971B-1
Sequence 1, Application US/07641971B
Patent No. 5236706

GENERAL INFORMATION:
APPLICANT: Debre, Patrice
APPLICANT: Mossalayi, Mohammed D
TITLE OF INVENTION: A PHARMACEUTICAL PREPARATION FOR THE
TITLE OF INVENTION: MATURATION OF PROTHYMOCYTES
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
STREET: 556 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/641,971B
FILING DATE: 19910116
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 90016254
FILING DATE: 24-JAN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Fishman, Irving M
REGISTRATION NUMBER: 30258
REFERENCE/DOCKET NUMBER: 4-17921/+/DEB

TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-277-4832
TELEFAX: 908-277-4306
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
CELL TYPE: Human B. Cells
CELL LINE: CHO cells transformed with pcAL8-BF-ND

Query Match 13.4%; Score 149; DB 1; Length 174;
US-07-641-971B-1

Query Match 13.4%; Score 149; DB 1; Length 174;
Best Local Similarity 30.2%; Pred. No. 5.7e-08;
Matches 55; Conservative 20; Mismatches 57; Indels 50; Gaps 11;

Best Local Similarity 30.2%; Pred. No. 5.7e-08;
Matches 55; Conservative 20; Mismatches 57; Indels 50; Gaps 11;

RESULT 9
US-07-781-248A-1
Sequence 1, Application US/07781248A
Patent No. 5246699

GENERAL INFORMATION:
APPLICANT: Debre, Patrice
APPLICANT: Mossalayi, Mohammed D
TITLE OF INVENTION: MATURATION OF HEMATOPOIETIC CELLS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
STREET: 556 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/781,248A
FILING DATE: 19911230
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: GB 90103565
FILING DATE: 09-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ikeler, Barbara J.
REGISTRATION NUMBER: 36,170
REFERENCE/DOCKET NUMBER: 4-18065/A/DEB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-277-3368
TELEFAX: 908-277-4306
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: AMINO ACID
TOPOLOGY: Linear
MOLECULE TYPE: Peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
CELL TYPE: Human B. Cells
CELL LINE: CHO cells transformed with pcAL8-BF-ND

Query Match 13.4%; Score 149; DB 1; Length 174;

Qy 17 ORPCYKVIYFHDTSRRLNFEAEAKEACRRDGGOLVSISEDEQKLIEKFIENLLPSDGDFW 76
 Db 24 ORKCY---YFGKGTKQ--WVHARYACDDMBEQOLVSIHSPEEQDFLTKH---ASHTGSW 73

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGS--EVCVVMMHQPSAPAGI 134
 Db 74 IGLRNLDLKG---FIWVDGSHVYDYSNWAPEPTSRSGEDCVMM---RGS 118

RESULT 11
 US-08-365-103B-8
 ; Sequence 8, Application US/08365103B
 ; Patent No. 5766943
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynch, Richard G
 ; APPLICANT: Nunez, Raphael D.
 ; APPLICANT: Yodoi, Jungi
 ; TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
 ; NUMBER OF SEQUENCES: 14
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
 ; STREET: 801 Grand Ave. Suite 3200
 ; CITY: Des Moines
 ; STATE: Iowa
 ; COUNTRY: United States
 ; ZIP: 50309
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/365,103B
 ; FILING DATE: 28-DEC-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Nebel, Heidi S.
 ; REGISTRATION NUMBER: 37,719
 ; REFERENCE/DOCKET NUMBER: Uirf N5-24
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (515) 288-3667
 ; TELEFAX: (515) 288-1338
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 321 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-365-103B-8

Query Match 13.4%; Score 149; DB 1; Length 321;
 Best Local Similarity 30.2%; Pred. No. 1.3e-07;
 Matches 55; Conservative 20; Mismatches 57; Indels 50; Gaps 11;

Qy 17 QRPCYKVIYFHDTSRRLNFEAEAKEACRRDGGOLVSISEDEQKLIEKFIENLLPSDGDFW 76
 Db 171 ORKCY---YFGKGTKQ--WVHARYACDDMEQOLVSIHSPEEQDFLTKH---ASHTGSW 220

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGS--EVCVVMMHQPSAPAGI 134
 Db 221 IGLRNLDLKG---FIWVDGSHVYDYSNWAPEPTSRSGEDCVMM---RGS 265

Query Match 13.4%; Score 149; DB 1; Length 320;
 Best Local Similarity 30.2%; Pred. No. 1.3e-07;
 Matches 55; Conservative 20; Mismatches 57; Indels 50; Gaps 11;

Qy 17 QRPCYKVIYFHDTSRRLNFEAEAKEACRRDGGOLVSISEDEQKLIEKFIENLLPSDGDFW 76
 Db 170 ORKCY---YFGKGTKQ--WVHARYACDDMEQOLVSIHSPEEQDFLTKH---ASHTGSW 219

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGS--EVCVVMMHQPSAPAGI 134
 Db 220 IGLRNLDLKG---FIWVDGSHVYDYSNWAPEPTSRSGEDCVMM---RGS 264

Query Match 13.4%; Score 149; DB 1; Length 320;
 Best Local Similarity 30.2%; Pred. No. 1.3e-07;
 Matches 55; Conservative 20; Mismatches 57; Indels 50; Gaps 11;

Qy 17 QRPCYKVIYFHDTSRRLNFEAEAKEACRRDGGOLVSISEDEQKLIEKFIENLLPSDGDFW 76
 Db 266 G----RWNDAFCDRKLGAWVC---DRLATCTPPASEGSAESMGPDSPRDPDGRLLPTPS 316

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGS--EVCVVMMHQPSAPAGI 134
 Db 317 AP 318

RESULT 12
 US-08-751-305-2
 ; Sequence 2, Application US/08751305

Patent No. 5965439
 GENERAL INFORMATION:
 APPLICANT: Tenner et al., Andrea J.
 TITLE OF INVENTION: HOST DEFENSE ENHANCEMENT
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 4225 Executive Square, Suite 1400
 CITY: La Jolla
 STATE: CA
 COUNTRY: USA
 ZIP: 92037
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/751,305
 FILING DATE: 18-NOV-1996
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Wetherell, Jr., John R.
 REGISTRATION NUMBER: 31,678
 REFERENCE/DOCKET NUMBER: 07306/012001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 619/678-5070
 TELEFAX: 619/678-5099
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 652 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-751-305-2

Query Match 13.3%; Score 148.5%; DB 2; Length 652;
 Best Local Similarity 24.4%; Pred. No. 3.9e-07;
 Matches 50; Conservative 37; Mismatches 75; Indels 43; Gaps 9;

5 LLSGQP-----VCRGGTQRPCYKVIYFHDTSTRRLNFEAKEACRRDGGQLVSI
 Db 13 LLIQPGAGTGADTEAVVCVG---TACYTA---HSGKLAAEAQNHCNQNGNLATVK 64

Qy 54 SEDEQKLIEKFIELNLLPSD-----GDFWIGLRRREEQKQSNSTACQDLIYAWT-DGSISQ 105
 Db 65 SKEEAQHVQRVLAQQLRREALTARMSKFWIGLQREKGKCLDPSLPLKGFSWVGGEDTP 124

Qy 106 FRNWYVD-EPSCGSEVCVVM---YHQPSA PAGIGGPPMFQWNDDRCNMKNN-----FI 154
 Db 125 YSNWHKELRNSCISKRCVSLLDLSQPLPNRLP----KWSSEGPGCGSPGSPGSNIEGFV 179

Qy 155 CKYSDKEPKAVPSREAEGEETELTP 179
 Db 180 CKFSEFKGMCRPLAAGGGPQVTTTP 204

RESULT 13 US-09-517-605-2
 Sequence 2, Application US/09517605
 Patent No. 6391567
 GENERAL INFORMATION:
 APPLICANT: Littman, Dan R.
 APPLICANT: Kwon, Douglas S.
 APPLICANT: van Kooyk, Yvette
 APPLICANT: Geijtenbeck, Theo
 TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
 FILE REFERENCE: 1049-1-017
 CURRENT APPLICATION NUMBER: US/09/517,605
 CURRENT FILING DATE: 2000-03-02
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2 ; LENGTH: 404 ;
; TYPE: PRT ; ORGANISM: Homo sapiens ; US-09-517-605-2 ;
; CORRESPONDENCE ADDRESS: Query Match 13.0%; Score 144.5%; DB 4; Length 404;
; Best Local Similarity 29.7%; Pred. No. 5.3e-07; Matches 47; Conservative 23; Mismatches 47; Indels 41; Gaps 10;

Qy 25 YFHDTSRRLNFEAKEACRRDGGQLVSVIESDEQKLIEKFIELNLLPSDGDFWIGLRRREE 84
 Db 268 YFMSNSQR-NWHDSTIACKEVGAQLVVIKSAAEQNFLQLQSSR--SNRFTWMGLSDLNQ 323

Qy 85 KQSNSTACQDLIYAWTDS--ISQFRN-WYVDEP-SCGSEVCVVMYHQPSA PAGIGGPPMF 140
 Db 324 -EGTWQWVGDGSPPLSPFQYWNRGEPPNNVGEEDCAEF-----SGNG----- 363

Qy 141 QWNDDRCNMKNNFICKYS-----DEK-----PAVPS 166
 Db 364 -WNDDKCNLAKFWICKKSAASCSRDEEQFLSPAPATPN 400

RESULT 14 US-08-513-278-4
; Sequence 4, Application US/08513278
; Patent No. 5840844
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,278
; FILING DATE: 10-AUG-1995
; CLASSIFICATION: 5530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreyer, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 565D1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-513-278-4

```

Query Match 12.9%; Score 143.5; DB 2; Length 372;
Best Local Similarity 26.5%; Pred. No. 6e-07;
Matches 41; Conservative 37; Mismatches 50; Indels 27; Gaps 8;

Qy 26 FHDTSRRLNFEAAKEACRRDGQLVSISSDEQKLIKEKFIELNL-P-SDGDFWIGLRRREE 84
Db 41 YHYSEKPMNWEARKFKQNYTDLVAIQNKRE--IE-YLENTLPKSPYYWIGRK-- 93

Qy 85 KQSNSTACQDLYAW--TDGSIS-QFRNWWVDEPS--CGSEVCVVMYHQPSAPAGIGGPM 139
Db 94 -----IGKMMWTWGTNKTKEAENWGAEPNNKSKEDCVEIYIKRERDSG----- 140

Qy 140 FQWNRDRCNMKNNFICKYSDEKPAVPSREAEGET 174
Db 141 -KWNDDACHKRKAALCYTASCPGSCNGRGEVCET 174

```

```

RESULT 15
5514582-4
; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO:4:
; LENGTH: 372
5514582-4

```

```

Query Match 12.9%; Score 143.5; DB 6; Length 372;
Best Local Similarity 26.5%; Pred. No. 6e-07;
Matches 41; Conservative 37; Mismatches 50; Indels 27; Gaps 8;

Qy 26 FHDTSRRLNFEAAKEACRRDGQLVSISSDEQKLIKEKFIELNL-P-SDGDFWIGLRRREE 84
Db 41 YHYSEKPMNWEARKFKQNYTDLVAIQNKRE--IE-YLENTLPKSPYYWIGRK-- 93

Qy 85 KQSNSTACQDLYAW--TDGSIS-QFRNWWVDEPS--CGSEVCVVMYHQPSAPAGIGGPM 139
Db 94 -----IGKMMWTWGTNKTKEAENWGAEPNNKSKEDCVEIYIKRERDSG----- 140

Qy 140 FQWNRDRCNMKNNFICKYSDEKPAVPSREAEGET 174
Db 141 -KWNDDACHKRKAALCYTASCPGSCNGRGEVCET 174

```

Search completed: May 20, 2003, 17:07:16
 Job time: 17 secs

GenCore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:06:10 : Search time 56 Seconds
 (without alignments)
 364.814 Million cell updates/sec

Title: US-09-887-855-5
 Perfect score: 1115
 Sequence: 1 ATGRLLSGQPVCRGQTQRPC.....EEDAKTKFESREAAALNAY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubcaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubcaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubcaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubcaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubcaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubcaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubcaa/PICTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubcaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubcaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubcaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubcaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubcaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubcaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubcaa/US60_PUBCOMB.pep:*
```

ALIGNMENTS

RESULT 1
 US-09-887-855-5
 ; Sequence 5, Application US/09887855
 ; Patent No. US20020058310A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Immunex Corporation
 ; INVENTOR: Anderson, Dirk M
 ; TITLE OF INVENTION: LECTIN SS3939 DNA AND POLYPEPTIDES
 ; FILE REFERENCE: 2883-US
 ; CURRENT APPLICATION NUMBER: US/09/887,855
 ; CURRENT FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 9
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 206
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-887-855-5

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1115	100.0	206	10 US-09-887-855-5	Sequence 5, App1
2	1115	100.0	374	9 US-10-149-819-15	Sequence 15, App1
3	1115	100.0	374	10 US-09-887-855-2	Sequence 2, App1
4	1101	98.7	382	9 US-09-905-291A-137	Sequence 137, App1
5	1101	98.7	382	9 US-09-902-853-137	Sequence 137, App1
6	1101	98.7	382	9 US-09-907-824-137	Sequence 137, App1
7	1101	98.7	382	9 US-09-907-841-137	Sequence 137, App1
8	1101	98.7	382	9 US-09-904-011-137	Sequence 137, App1
9	1101	98.7	382	9 US-10-174-590-20	Sequence 20, App1
10	1101	98.7	382	9 US-10-176-758-20	Sequence 20, App1
11	1101	98.7	382	9 US-10-175-737-20	Sequence 20, App1
12	1101	98.7	382	9 US-09-906-742-137	Sequence 137, App1
13	1101	98.7	382	9 US-10-173-706-20	Sequence 20, App1
14	1101	98.7	382	9 US-10-175-738-20	Sequence 20, App1
15	1101	98.7	382	9 US-10-175-752-20	Sequence 20, App1
16	1101	98.7	382	9 US-10-176-482-20	Sequence 20, App1
17	1101	98.7	382	9 US-10-176-757-20	Sequence 20, App1
18	1101	98.7	382	9 US-10-176-913-20	Sequence 20, App1
19	1101	98.7	382	9 US-10-180-552-20	Sequence 20, App1

RESULT 2
 US-10-149-819-15

Sequence 15, Application US/10149819

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: AZIMZAI, Yalda

APPLICANT: TANG, Y. Tom

APPLICANT: PATTERSON, Chandra

APPLICANT: BAUGHN, Mariah R.

APPLICANT: LU, Dyung Aina M.

APPLICANT: SHAH, Purvi

APPLICANT: LAL, Preeti

APPLICANT: AU-YOUNG, Janice

APPLICANT: BURFORD, Neil

TITLE OF INVENTION: EXTRACELLULAR MATRIX AND CELL ADHESION MOLECULES

FILE REFERENCE: PF-0760 PCT

CURRENT APPLICATION NUMBER: US/10/149,819

CURRENT FILING DATE: 2002-06-10

PRIOR APPLICATION NUMBER: 60/172,852; 60/172,354

PRIOR FILING DATE: 1999-12-10; 1999-12-16

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PERL Program

SEQ ID NO 15

LENGTH: 374

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

OTHER INFORMATION: Incyte ID NO. US20030044913A1 3143411CD1

US-10-149-819-15

Query Match 100.0%; Score 1115; DB 9; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.4e-102;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60

Db 22 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy 61 IEKFIELLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120

Db 82 IEKFIELLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy 121 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 180

Db 142 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 201

Qy 181 LPEETQEDAKKTFKESREAAALNAY 206

Db 202 LPEETQEDAKKTFKESREAAALNAY 227

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 60

Db 22 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRRDGGQLVSIESEDEQKL 81

Qy 61 IEKFIELLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 120

Db 82 IEKFIELLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPSCGSEV 141

Qy 121 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 180

Db 142 CVMYHQPSAPAGIGGPPMFQWNDDRCNMKNFICKYSDEKPAVPSREAEGEETELTPV 201

Qy 181 LPEETQEDAKKTFKESREAAALNAY 206

Db 202 LPEETQEDAKKTFKESREAAALNAY 227

RESULT 4

US-09-905-291A-137

; Sequence 137, Application US/09905291A

; Patent No. US20020160374A1

; GENERAL INFORMATION:

; APPLICANT: Genentech, Inc.

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Ferrara, Napoleone

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, Christopher J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillian, Kenneth, J.

; APPLICANT: Kijavin, Ivar J.

; APPLICANT: Mather, Jennie P.

; APPLICANT: Pan, James

; APPLICANT: Paonil, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Steward, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William, I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: 10466-14

; CURRENT APPLICATION NUMBER: US/09/905,291A

; CURRENT FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: PCT/US00/04414

; PRIOR FILING DATE: 2000-02-22

; PRIOR APPLICATION NUMBER: US 60/143,048

; PRIOR FILING DATE: 1999-07-07

; PRIOR APPLICATION NUMBER: US 60/145,698

; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222

; PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: PCT/US99/20594

; PRIOR FILING DATE: 1999-09-08

; PRIOR APPLICATION NUMBER: PCT/US99/20944

; PRIOR FILING DATE: 1999-09-13

; PRIOR APPLICATION NUMBER: PCT/US99/21090

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/21547

; PRIOR FILING DATE: 1999-09-15

; PRIOR APPLICATION NUMBER: PCT/US99/23089

; PRIOR FILING DATE: 1999-10-05

; PRIOR APPLICATION NUMBER: PCT/US99/28214

Query Match 100.0%; Score 1115; DB 10; Length 374;

Best Local Similarity 100.0%; Pred. No. 1.4e-102;

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 112
 Db 82 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 141

Qy 113 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDERPAVPSREAGE 172
 Db 142 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201

Qy 173 ETELTPVLPTEETQEEDAKTFKESREALNLAY 206
 Db 202 ETELTPVLPTEETQEEDAKTFKESREALNLAY 235

RESULT 5
 US-09-902-853-137
 Sequence 137, Application US/09902853
 Publication No. US20020192659A1
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, A.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: KJavin, Kenneth, J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.

RESULT 6
 US-09-907-824-137
 Sequence 137, Application US/09907824
 Publication No. US20020197671A1
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 112
 Db 82 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 141

Qy 113 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDERPAVPSREAGE 172
 Db 142 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201

Qy 173 ETELTPVLPTEETQEEDAKTFKESREALNLAY 206
 Db 202 ETELTPVLPTEETQEEDAKTFKESREALNLAY 235

RESULT 5
 US-09-902-853-137
 Sequence 137, Application US/09902853
 Publication No. US20020192659A1
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.
 APPLICANT: Ashkenazi, Avi
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, A.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: KJavin, Kenneth, J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 112
 Db 82 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNINSTACQDLYAWTDGSISQFRNWYD 141

Qy 113 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDERPAVPSREAGE 172
 Db 142 EPSCGSEEVVVMYHQPSAPAGIGGPMFQWNDRCNMKNNFICKYSDEKPAVPSREAGE 201

Qy 173 ETELTPVLPTEETQEEDAKTFKESREALNLAY 206
 Db 202 ETELTPVLPTEETQEEDAKTFKESREALNLAY 235

APPLICANT: Ashkenazi, Avi
 Botstein, David
 Desnoyers, Luc
 Applicant: Eaton, Dan L.
 Applicant: Ferrara, Napoleone
 Applicant: Filvaroff, Ellen
 Applicant: Fong, Sherman
 Applicant: Gao, Wei-Qiang
 Applicant: Gerber, Hanspeter
 Applicant: Gerritsen, Mary E.
 Applicant: Goddard, A.
 Applicant: Godowski, Paul J.
 Applicant: Grimaldi, Christopher J.
 Applicant: Gurley, Austin L.
 Applicant: Hillian, Kenneth J.
 Applicant: Klijavin, Ivar J.
 Applicant: Mather, Jennie P.
 Applicant: Pan, James
 Applicant: Paoni, Nicholas F.
 Applicant: Roy, Margaret Ann
 Applicant: Stewart, Timothy A.
 Applicant: Tunas, Daniel
 Applicant: Williams, P. Mickey
 Applicant: Wood, William J.
 TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 TITLE OF INVENTION: Acids Encoding the Same
 FILE REFERENCE: 10466-14
 CURRENT APPLICATION NUMBER: US/09/907,824
 CURRENT FILING DATE: 2001-07-17
 PRIOR APPLICATION NUMBER: 09/665,350
 PRIOR FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: US 60/143,048
 PRIOR FILING DATE: 1999-07-07
 PRIOR APPLICATION NUMBER: US 60/145,698
 PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: US 60/146,222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: PCT/US99/20594
 PRIOR FILING DATE: 1999-09-08
 PRIOR APPLICATION NUMBER: PCT/US99/20944
 PRIOR FILING DATE: 1999-09-13
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/21547
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/23089
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 137
 LENGTH: 382
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-907-824-137
 Query Match 98.7%; Score 1101; DB 9; Length 382;

Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIVFHDTSRRLNFEEAKEACRRDGGQLVSI 52
 Db 22 ATGRLLSASDDLRLGGQPVCRGGTQRPCYKVIVFHDTSRRLNFEEAKEACRRDGGQLVSI 81
 Qy 53 ESEDEQKLIEKFENLLPSPGDWFIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 112
 Db 82 ESEDEQKLIEKFENLLPSPGDWFIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 141
 Qy 113 EPSCGSECVVMMYHOPSPAPAGIGGPMFQWNDDRCNMKNNFICKYSDKEPKAVPSREAEQE 172
 Db 142 EPSCGSECVVMMYHOPSPAPAGIGGPMFQWNDDRCNMKNNFICKYSDKEPKAVPSREAEQE 201
 Qy 173 ETELTPVLPETQEEDAKTFKESREALNLAY 206
 Db 202 ETELTPVLPETQEEDAKTFKESREALNLAY 235

RESULT 7
 US-09-907-841-137
 ; Sequence 137, Application US/09907841
 ; Publication No. US20020198366A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, A.
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurley, Austin L.
 ; APPLICANT: Hillian, Kenneth J.
 ; APPLICANT: Klijavin, Ivar J.
 ; APPLICANT: Mather, Jennie P.
 ; APPLICANT: Pan, James
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tunas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William J.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: 10466-14
 ; CURRENT APPLICATION NUMBER: US/09/907,824
 ; CURRENT FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: 09/665,350
 ; PRIOR FILING DATE: 2000-09-18
 ; PRIOR APPLICATION NUMBER: PCT/US00/04414
 ; PRIOR FILING DATE: 2000-02-22
 ; PRIOR APPLICATION NUMBER: US 60/143,048
 ; PRIOR FILING DATE: 1999-07-07
 ; PRIOR APPLICATION NUMBER: US 60/145,698
 ; PRIOR FILING DATE: 1999-07-26
 ; PRIOR APPLICATION NUMBER: US 60/146,222
 ; PRIOR FILING DATE: 1999-07-28
 ; PRIOR APPLICATION NUMBER: PCT/US99/20594
 ; PRIOR FILING DATE: 1999-09-08
 ; PRIOR APPLICATION NUMBER: PCT/US99/20944
 ; PRIOR FILING DATE: 1999-09-13
 ; PRIOR APPLICATION NUMBER: PCT/US99/21090
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/21547
 ; PRIOR FILING DATE: 1999-09-15
 ; PRIOR APPLICATION NUMBER: PCT/US99/23089
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: PCT/US99/28214
 ; PRIOR FILING DATE: 1999-11-29
 ; PRIOR APPLICATION NUMBER: PCT/US99/28313
 ; PRIOR FILING DATE: 1999-11-30
 ; PRIOR APPLICATION NUMBER: PCT/US99/28564
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/28565
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30095
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: PCT/US99/30911
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US99/30999
 ; PRIOR FILING DATE: 1999-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US00/00219
 ; PRIOR FILING DATE: 2000-01-05
 ; NUMBER OF SEQ ID NOS: 423
 ; SEQ ID NO 137
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-09-907-824-137
 ; Query Match 98.7%; Score 1101; DB 9; Length 382;

PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 137
 LENGTH: 382
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-907-841-137

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRQDGQLVSI 52
 Db 22 ATGRLLSASDLDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRQDGQLVSI 81
 Qy 53 ESEDEQKLIEKFIEENLLPSDGFDFWIGLRLRREEKQSNSTACQDLYAWTDGSISQFRNWVD 112
 Db 82 ESEDEQKLIEKFIEENLLPSDGFDFWIGLRLRREEKQSNSTACQDLYAWTDGSISQFRNWVD 141
 Qy 113 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAGE 172
 Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAGE 201
 Qy 173 ETELTPVLPTEETQEEAKKTFKESREALNLAY 206
 Db 202 ETELTPVLPTEETQEEAKKTFKESREALNLAY 235

RESULT 8
 US-09-904-011-137
 Sequence 137, Application US/09904011
 Publication No. US20030003530A1
 GENERAL INFORMATION:
 APPLICANT: Genentech, Inc.

APPLICANT: Ashkenazi, Avi
 APPLICANT: Botstein, David
 APPLICANT: Desnoyers, Luc
 APPLICANT: Eaton, Dan L.
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Fong, Sherman
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillian, Kenneth, J.
 APPLICANT: Klijavin, Ivar J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.

TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acid Encoding the Same
 FILE REFERENCE: 10466-14

CURRENT APPLICATION NUMBER: US/09/904,011
 CURRENT FILING DATE: 2001-07-11
 PRIOR APPLICATION NUMBER: 09/665, 350
 PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: US 60/143, 048
 PRIOR FILING DATE: 1999-07-07

PRIOR APPLICATION NUMBER: US 60/145, 698

PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: US 60/146, 222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: PCT/US99/20594
 PRIOR FILING DATE: 1999-09-08
 PRIOR APPLICATION NUMBER: PCT/US99/20944
 PRIOR FILING DATE: 1999-09-13
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/21547
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/23089
 PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: 1999-12-16
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 137
 LENGTH: 382
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-904-011-137

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;
 Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRQDGQLVSI 52
 Db 22 ATGRLLSASDLDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRQDGQLVSI 81
 Qy 53 ESEDEQKLIEKFIEENLLPSDGFDFWIGLRLRREEKQSNSTACQDLYAWTDGSISQFRNWVD 112
 Db 82 ESEDEQKLIEKFIEENLLPSDGFDFWIGLRLRREEKQSNSTACQDLYAWTDGSISQFRNWVD 141
 Qy 113 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAGE 172
 Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAGE 201
 Qy 173 ETELTPVLPTEETQEEAKKTFKESREALNLAY 206
 Db 202 ETELTPVLPTEETQEEAKKTFKESREALNLAY 235

RESULT 9
 US-10-174-590-20
 Sequence 20, Application US/10174590
 Publication No. US20030008352A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin P.
 APPLICANT: Chen, Jian
 APPLICANT: Desnoyers, Luc
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
 FILE REFERENCE: P3430R1C42
 CURRENT APPLICATION NUMBER: US/10/174, 590
 CURRENT FILING DATE: 2002-06-18
 PRIORITY APPLICATION removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 612
 SEQ ID NO 20
 LENGTH: 382
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-174-590-20

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
 Db 22 ATGRLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141
 Db 82 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 141

Qy 113 EPSCGSEVCVVMYHQPSAPAGIGGPPMFQWNDDRCMKNNNFICKYSDEKPAVPSREAEGE 172
 Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPPMFQWNDDRCMKNNNFICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTTPVLPETQEEAKKTFKESREALNLAY 206
 Db 202 ETELTTPVLPETQEEAKKTFKESREALNLAY 235

RESULT 10
 US-10-176-758-20
 ; Sequence 20, Application.US/10176758
 ; Publication No. US20030008353A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3430R1C50
 ; CURRENT APPLICATION NUMBER: US/10/175, 737
 ; CURRENT FILING DATE: 2002-06-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 20
 ; LENGTH: 382
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-175-737-20

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
 Db 22 ATGRLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112
 Db 82 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112

Qy 113 EPSCGSEVCVVMYHQPSAPAGIGGPPMFQWNDDRCMKNNNFICKYSDEKPAVPSREAEGE 141
 Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPPMFQWNDDRCMKNNNFICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTTPVLPETQEEAKKTFKESREALNLAY 206
 Db 202 ETELTTPVLPETQEEAKKTFKESREALNLAY 235

RESULT 12
 US-09-906-742-137
 ; Sequence 137, Application US/0906742
 ; Publication No. US20030023054A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Genentech, Inc.
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Ferrara, Napoleone
 ; APPLICANT: Filvaroff, Ellen

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 52
 Db 22 ATGRLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSI 81

Qy 53 ESEDEQKLIKEKFIEENLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVD 112

APPLICANT: Fong, Sherman QY 53 ESEDEQKLIKFIELNLLPSDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 112
 APPLICANT: Gao, Wei-Qiang Db 82 ESEDEQKLIKFIELNLLPSDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 141
 APPLICANT: Gerber, Hanspeter
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, A.
 APPLICANT: Godowski, Paul J.
 APPLICANT: Grimaldi, Christopher J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Hillan, Kenneth, J.
 APPLICANT: Kijavin, Ivar J.
 APPLICANT: Mather, Jennie P.
 APPLICANT: Pan, James
 APPLICANT: Paoni, Nicholas F.
 APPLICANT: Roy, Margaret Ann
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Williams, P. Mickey
 APPLICANT: Wood, William, I.
 APPLICANT: Wood, William I.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin

TITLE OF INVENTION: Secreted and transmembrane Polypeptides and Nucleic
 FILE REFERENCE: 10466-14
 CURRENT APPLICATION NUMBER: US/09/906,742
 CURRENT FILING DATE: 2001-07-16
 PRIOR APPLICATION NUMBER: 09/665,350
 PRIOR FILING DATE: 2000-09-18
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: 2000-02-22
 PRIOR APPLICATION NUMBER: US 60/143,048
 PRIOR FILING DATE: 1999-07-07
 PRIOR APPLICATION NUMBER: US 60/145,698
 PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: US 60/146,222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: PCT/US99/20594
 PRIOR FILING DATE: 1999-09-08
 PRIOR APPLICATION NUMBER: PCT/US99/20944
 PRIOR FILING DATE: 1999-09-13
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/21547
 PRIOR FILING DATE: 1999-09-15
 PRIOR APPLICATION NUMBER: PCT/US99/23089
 PRIOR FILING DATE: 1999-10-05
 PRIOR APPLICATION NUMBER: PCT/US99/28214
 PRIOR FILING DATE: 1999-11-29
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: 1999-11-30
 PRIOR APPLICATION NUMBER: PCT/US99/28564
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/28565
 PRIOR FILING DATE: 1999-12-02
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR APPLICATION NUMBER: PCT/US99/30911
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US99/30999
 PRIOR FILING DATE: 1999-12-20
 PRIOR APPLICATION NUMBER: PCT/US00/00219
 PRIOR FILING DATE: 2000-01-05
 NUMBER OF SEQ ID NOS: 423
 SEQ ID NO 137
 LENGTH: 382
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-906-742-137

RESULT 13
 US-10-173-706-20
 ; Sequence 20, Application US/10173706
 ; Publication No. US20030022293A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSI 81

QY 53 ESEDEQKLIKFIELNLLPSDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 112
 Db 82 ESEDEQKLIKFIELNLLPSDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVD 141

QY 113 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDRCNMNNFICKYSDEKPAVPSREAEGE 172
 Db 142 EPSCGSEVCVVMYHQPSAPAGIGGPMFQWNDRCNMNNFICKYSDEKPAVPSREAEGE 201

QY 173 ETELTTTPVLPTEETQEEDAKKTFKESREAAALNAY 206
 Db 202 ETELTTTPVLPTEETQEEDAKKTFKESREAAALNAY 235

RESULT 14
 US-10-175-738-20
 ; Sequence 20, Application US/10175738
 ; Publication No. US20030022294A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 1 ATGRLLS-----GQPVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSI 81

APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Pan, James
 APPLICANT: Smith, Victoria
 APPLICANT: Watanabe, Colin K.
 APPLICANT: Wood, William I.
 APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C45

CURRENT APPLICATION NUMBER: US/10/175,738

CURRENT FILING DATE: 2002-06-19

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 20

LENGTH: 382

TYPE: PRT

ORGANISM: Homo Sapien

US-10-175-738-20

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Qy 1 ATGRLLS-----GQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGQLVSI 52
 Db 22 ATGRLLSASDLDLRRGGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGQLVSI 81

Qy 53 ESEDEQKLIKEKFIENLLPSDGFHWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWyVD 112
 Db 82 ESEDEQKLIKEKFIENLLPSDGFHWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWyVD 141

Qy 113 EPSCGSEVCVVMMYHQPSAPAGIGGPMFQWINDDRCNMKNNFICKYSDEKPAVPSREAEGE 172
 Db 142 EPSCGSEVCVVMMYHQPSAPAGIGGPMFQWINDDRCNMKNNFICKYSDEKPAVPSREAEGE 201

Qy 173 ETELTPVLPTEETQEDAKKTFKESREAAALNAY 206
 Db 202 ETELTPVLPTEETQEDAKKTFKESREAAALNAY 235

RESULT 15
 US-10-175-752-20

; Sequence 20, Application US/10175752
 ; Publication No. US20030022295A1
 ; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C60

CURRENT APPLICATION NUMBER: US/10/175,752

CURRENT FILING DATE: 2002-06-19

Prior application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 20

LENGTH: 382

TYPE: PRT

ORGANISM: Homo Sapien

Query Match 98.7%; Score 1101; DB 9; Length 382;
 Best Local Similarity 96.3%; Pred. No. 3.6e-101;
 Matches 206; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

Copyright (c) 1993 - 2003 Compugen Ltd. GenCore version 5.1.4_p5_4578

NCBInfo protein - protein search, using SW model

Run on: May 20, 2003, 17:03:25 ; Search time 44 seconds

118-09-887-855-5
[t]e;

Perfect score: 1115
Title: 03-09-88/-833-3

bioRxiv preprint doi: <https://doi.org/10.1101/2022.05.10.488082>; this version posted May 10, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under a [aCC-BY-ND 4.0 International license](https://creativecommons.org/licenses/by-nd/4.0/).

searched: 383224 segs: 96134422 residues

total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Rating Error 45.250000000000004

Database : BIB 73:*

```
1:  pir1:=*  
2:  pir2:=*  
3:  Pir3:=*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	182	16.3	1456	1	A36563		mannose receptor P
2	178.5	16.0	1455	1	A48925		mannose receptor P
3	177	15.9	1268	2	S52781		neurocan - mouse
4	174.5	15.7	1643	2	T14274		versican precursor
5	174.5	15.7	3381	2	T42389		versican precursor
6	174	15.6	1257	2	S28764		neurocan precursor
7	174	15.6	2397	1	A55535		versican precursor
8	174	15.6	2409	1	A60979		versican precursor
9	171	15.3	3562	2	A47171		chondroitin sulfat
10	170.5	15.3	1479	2	T42710		mannose receptor,
11	158.5	14.2	1340	2	A39808		proteoglycan core
12	158.5	14.2	2327	2	T42630		aggreccan - bovine
13	158.5	14.2	2415	1	A39086		aggreccan precursor
14	154.5	13.9	612	2	B42755		E-selectin precurs
15	153.5	13.8	2124	2	A28452		proteoglycan core
16	152	13.6	912	2	A54423		brevican precursor
17	149.5	13.4	459	2	T24425		hypothetical prote
18	149	13.4	321	1	LNUER		IgE Fc receptor II
19	148.5	13.3	330	2	T46256		brevican - human (
20	148.5	13.3	2132	1	A55182		aggreccan precursor
21	148	13.3	253	2	E89130		protein F52E1.2 [1
22	147	13.2	883	2	S57653		brevican precursor
23	146.5	13.1	162	1	LNRC1		lectin BRA3-1 prec
24	146	13.1	2109	1	I50421		aggreccan precursor
25	145.5	13.0	742	2	JC7595		scavenger receptor
26	145	13.0	883	2	S49126		brevican precursor
27	144.5	13.0	173	2	S10548		lectin - barnacle
28	144.5	13.0	372	2	S23936		L-selectin precurs
29	144.5	13.0	404	2	A46274		HIV gp120 binding

30	143.5	12.9	JC4329
31	143.5	12.9	372 1 A32375
32	143.5	12.9	463 2 T26655
33	142.5	12.8	131 2 JC5058
34	142	12.7	1487 2 S48719
35	141.5	12.7	331 1 LNMSER
36	140.5	12.6	162 1 LNRC3
37	140	12.6	370 2 S22124
38	139	12.5	248 1 LNHUPS
39	139	12.5	248 1 LNHUP6
40	139	12.5	248 1 LNHUP1
41	139	12.5	283 1 LNFHLS
42	138.5	12.4	152 2 JC4690
43	138.5	12.4	202 2 JC4031
44	138	12.4	280 2 T29200
45	137.5	12.3	309 1 S34198

ALIGNMENTS

RESULT 1
A36563
mannose receptor precursor - human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999
C;Accession: A36563; A60926; A44255; B44255; C44255; D44255; E44255; F44255; G44255;
R;Taylor, M.E.; Conary, J.T.; Lennartz, M.R.; Stahl, P.D.; Drickamer, K.
J. Biol. Chem. 265, 12156-12162, 1990
A;Title: Primary structure of the mannose receptor contains multiple motifs resembling
A;Reference number: A36563; MUID:90324192; PMID:2373685

A; Molecule type: mRNA
A; Residues: 1-1456 <TAY>
A; Cross-references: GB:J05550; NID:9188675; PID:AAA59868.1; PMID:9188676
A; Note: parts of this sequence, including the amino end of the mature protein, were
R; Ezekowitz, R.A.B.; Sastry, K.; Bailly, P.; Warner, A.
J. Exp. Med. 172, 1785-1794, 1990
A; Title: Molecular characterization of the human macrophage mannose receptor: demo
A; Reference number: A60926; MUID:91079783; PMID:2258707
A; Accession: A60926
A; Status: nucleic acid sequence not shown
A; Molecule type: mRNA
A; Residues: 1-1333, 'T', 1335-1456 <EE>
A; Cross-references: GB:X55635
A; Note: translation of the nucleotide sequence is incomplete
A; Note: in the authors' translation additional residues Pro-Glu-Ile are shown after
R; Kim, S.J.; Ruiz, N.; Bezouska, K.; Drickamer, K.
Genomics 14, 721-727, 1992
A; Title: Organization of the gene encoding the human macrophage mannose receptor (demo)
A; Reference number: A44255; MUID:93052405; PMID:1294118
A; Accession: A44255
A; Status: preliminary; nucleic acid sequence not shown; not compared with conceptual
A; Molecule type: DNA
A; Residues: 155-233, 'KSAL', 238-283; 346-428; 492-569; 631-714, 716-719; 783-820, 'N', 822
A; Note: sequence extracted from NCBI backbone (NCBIP:118415, NCBIP:118421, NCBIP:1
C; Genetics:
A; Gene: GDB:MRC1
A; Cross-references: GDB:133759; OMIM:153618
A; Map position: 10p13-10p13
C; Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type
C; Keywords: duplication; lectin; tandem repeat; transmembrane protein
F; 1-18/Domain: signal sequence #status predicted <SIG>
F; 168-209/Domain: fibronectin type II repeat homology <2F1>
F; 223-340/Domain: C-type lectin homology <LCH1>
F; 362-486/Domain: C-type lectin homology <LCH2>
F; 945-1079/Domain: C-type lectin homology <LCH3>

```

Query Match 16.3%; Score 182; DB 1; Length 1456;
Best Local Similarity 25.5%; Pred. No. 9.3e-08;
Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps
QY 21 XKVIXFHDTTSRRRLNFFFAKEAACRRDGGSIVSIESFDEOKLIFKEIENILPSDGDEWICIB 80

```


QY 76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWYVDEP---SCGSEVCVVMMYHQPSAP 131
 Db 1472 WIGL---NDKMFHEDFRWTGDSLQYENWRPNQPDSSFSTGEDCVVIIWHENG-- 1521

QY 132 AGIGGPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAGE--- 172
 Db 1522 -----QWNNDVPCNYHLYTCKKGTVACGQPPVENAKTFGMKPRYEINSLIRYHC 1572

QY 173 -----ETELTT---PVL---PEETQEEDEAKTFKESREALN 203
 Db 1573 KDGFIQRLPTIRCLGNRWAAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 1623

Query Match 15.6%; Score 174; DB 2; Length 1257;

Best Local Similarity 30.8%; Pred. No. 4e-07; Mismatches 49; Indels 32; Gaps 6;

Matches 44; Conservative 18; Mismatches 49; Indels 32; Gaps 6;

RESULT 5

T42389 versican precursor, splice form v0 - bovine
 N;Alternate names: chondroitin sulfate proteoglycan
 C;Species: Bos primigenius taurus (cattle)
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 05-May-2000
 C;Accession: T42389
 R;Schmalfieldt, M.; Dours-Zimmermann, M.T.; Winterhalter, K.H.; Zimmermann, D.R.
 J. Biol. Chem. 273, 15758-15764, 1998
 A;Title: Versican V2 is a major extracellular matrix component of the mature bovine brain
 A;Reference number: Z17954; MUID:98288320; PMID:9624174
 A;Accession: T42389
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-3381 <SCH>
 A;Cross-references: EMBL:AF060456; NID:93253299; PID:g3253300; PIDN: AAC24358.1
 C;Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type lectin
 C;Keywords: chondroitin sulfate proteoglycan; extracellular matrix; glycoprotein
 F;1-20/Domain: signal sequence #status Predicted <SIG>
 F;21-3381/Domain: versican, splice form v0 #status Predicted <MAT>
 F;57,331,352,817,965,1017,1333,1393,1437,1463,1653,1974,2045,2074,2103,2263,2290,2356,26
 Query Match 15.7%; Score 174.5; DB 2; Length 3381;
 Best Local Similarity 25.5%; Pred. No. 1.1e-06;
 Matches 59; Conservative 26; Mismatches 71; Indels 75; Gaps 11;

QY 17 QRPCYKVIYFHDTSRRLNFEEAKEACRQDGQLVSISEDEQKLIKEFIEENLPLPSDGFV 76
 Db 3162 QGQCYK--YF--AHRRTWDAEERCLQGAHLTSILSHEEQMFVNVRV---GHDYQ 3209

QY 76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWYVDEP---SCGSEVCVVMMYHQPSAP 131
 Db 3210 WIGL---NDKMFHEDFRWTGDSLQYENWRPNQPDSSFSTGEDCVVIIWHENG-- 3259

QY 132 AGIGGPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAGE--- 172
 Db 3260 -----QWNNDVPCNYHLYTCKKGTVACGQPPVENAKTFGMKPRYEINSLIRYHC 3310

QY 173 -----ETELTT---PVL---PEETQEEDEAKTFKESREALN 203
 Db 3311 KDGFIQRLPTIRCLGNRWAAMPKITCLNPSAYQRTYSKKYFKNSSSAKDN 3361

Query Match 15.6%; Score 174; DB 1; Length 2397;

Best Local Similarity 28.5%; Pred. No. 8.4e-07; Mismatches 55; Indels 40; Gaps 8;

Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

RESULT 6

S28764 neurocan precursor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 22-Nov-1993 #sequence_revision 01-Sep-1995 #text_change 04-Feb-2000
 C;Accession: S28764
 R;Rauch, U.; Karthikeyan, L.; Maurel, P.; Margolis, R.K.
 J. Biol. Chem. 267, 19536-19547, 1992
 A;Title: Cloning and primary structure of neurocan, a developmentally regulated, aggregated
 A;Reference number: S28764; MUID:92406907; PMID:1326557
 A;Accession: S28764
 A;Molecule type: mRNA
 A;Residues: 1-1257 <RAU>
 A;Cross-references: EMBL:M97161; NID:g205649; PIDN: AAC37679.1; PID:g205650
 C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG
 C;Keywords: chondroitin sulfate proteoglycan; glycoprotein

RESULT 8

A60979 versican precursor - human
N; Contains: chondroitin sulfate proteoglycan 2; chondroitin sulfate attachment *status predicted <GAG>
C; Species: Homo sapiens (man)
C; Date: 10-Sep-1999 *sequence_revision 10-Sep-1999 *text_change 19-Jan-2001
C; Accession: S06014; S43921; A60979; A30358; A29348; A45131; 154179
R; Zimmermann, D.R.; Ruoslahti, E.
EMBO J. 8, 2975-2981, 1989
A; Title: Multiple domains of the large fibroblast proteoglycan, versican.
A; Reference number: S06014; MUID:90059882; PMID:2583089
A; Accession: S06014
A; Molecule type: mRNA
A; Residues: 1-2409 <ZIM>
A; Cross-references: GB:X15998; NID:937662; PIDN:CAA34128.1; PID:937663
R; Yao, L.Y.; Moody, C.; Schoenherr, E.; Wight, T.N.; Sandell, L.J.
Matrix Biol. 14, 213-225, 1994
A; Title: Identification of the proteoglycan versican in aorta and smooth muscle cells by
A; Reference number: S43921; MUID:95005762; PMID:7921538
A; Accession: S43921
A; Molecule type: mRNA
A; Residues: 208-440; 1094-1385; 1910-2246 <YAO>
R; Bignami, A.; Lane, W.S.; Andrews, D.; Dahl, D.
Brain Res. Bull. 22, 67-70, 1989
A; Title: Structural similarity of hyaluronate binding proteins in brain and cartilage.
A; Reference number: A60979; MUID:89229983; PMID:2469524
A; Accession: A60979
A; Molecule type: protein
A; Residues: 171-210; 289-303 <BIG>
R; Perides, G.; Lane, W.S.; Andrews, D.; Dahl, D.; Bignami, A.
J. Biol. Chem. 264, 5981-5987, 1989
A; Title: Isolation and partial characterization of a glial hyaluronate-binding protein.
A; Reference number: A30358; MUID:89174663; PMID:2466833
A; Accession: A30358
A; Molecule type: protein
A; Residues: 24-50; 80-87, 'D', 89-119; 128-155; 167-218; 229-259, 'IR', 261-268; 277-283, 'G', 285-
R; Krusius, T.; Gehlsen, K.R.; Ruoslahti, E.
J. Biol. Chem. 262, 13120-13125, 1987
A; Title: A fibroblast chondroitin sulfate proteoglycan core protein contains lectin-like
A; Reference number: A29348; MUID:88007514; PMID:2820964
A; Accession: A29348
A; Molecule type: mRNA
A; Residues: 1725, 'V', 1727-2409 <KRU>
A; Cross-references: GB:J02814
R; Perides, G.; Rahemtulla, F.; Lane, W.S.; Asher, R.A.; Bignami, A.
J. Biol. Chem. 267, 23883-23887, 1992
A; Title: Isolation of a large aggregating proteoglycan from human brain.
A; Reference number: A45131; MUID:93054750; PMID:1429726
A; Contents: brain
A; Accession: A45131
A; Molecule type: protein
A; Residues: 21-22, 'X', 24-37 <PE2>
A; Experimental source: brain
A; Note: sequence extracted from NCBI backbone (NCBIP:118884)
R; Iozzo, R.V.; Naso, M.F.; Cannizzaro, L.A.; Wasmuth, J.J.; McPherson, J.D.
Genomics 14, 845-851, 1992
A; Title: Mapping of the versican proteoglycan gene (CSPG2) to the long arm of human chrc
A; Reference number: I54179; MUID:93122792; PMID:1478664
A; Accession: I54179
A; Status: translated from GB/EMBL/DDBJ
A; Molecule type: DNA
A; Residues: 251-347 <RES>
A; Cross-references: GB:S552488; NID:9263313; PIDN:AAB24878.1; PID:9263314
C; Genetics:

RESULT 8

F;265-346/Domain: link protein repeat homology <LNK2>
F;559-1654/Domain: chondroitin sulfate attachment *status predicted <GAG>
F;2106-2137/Domain: EGF homology <EG1>
F;2144-2175/Domain: EGF homology <EG2>
F;2182-2302/Domain: C-type lectin homology <LCH>
F;2309-2365/Domain: complement factor H repeat homology <FHD>
Query Match 15.6%; Score 174; DB 1; Length 2409;
Best Local Similarity 28.5%; Pred. No. 8.5e-07;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
Db 2190 QGQCYK--YF--AHRRTWDAEERECLQAHLTSILSHEEQMFVNRV----GHDYQ 2237
Qy 17 QRPCYKVIIYFHDTSRRLNFEAKEACRRDGGLVSISEDEQKLIEKFIEFLNLLPSDGF- 75
Db 2190 QGQCYK--YF--AHRRTWDAEERECLQAHLTSILSHEEQMFVNRV----GHDYQ 2237
Qy 7.6 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWFVDEP----SCGSEVCVVMYHOPSAF 131
Db 2238 WIGL----NDKMFEEHDFRWTGDSLQYENWRPNQPDSSFFSAGEDCVCVIIWHENG-- 2287
Qy 132 AGIGGPMQWNDDRCNMKNNFICKYS---DEKPAVPSREAEGE 172
Db 2288 -QWNDDVPCNYHLLTYTCKKGTVACGQPPVVENAKTFFGK 2323
RESULT 9
A47171
chondroitin sulfate proteoglycan PG-M core protein - chicken
C; Species: Gallus gallus (chicken)
C; Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 *text_change 21-Jul-2000
C; Accession: A47171
R; Shinomura, T.; Nishida, Y.; Ito, K.; Kimata, K.
J. Biol. Chem. 268, 14461-14469, 1993
A; Title: cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed du
A; Reference number: A47171; MUID:93300846; PMID:8314802
A; Accession: A47171
A; Status: preliminary
A; Molecule type: nucleic acid
A; Residues: 1-3562 <SH1>
A; Cross-references: GB:D13542; NID:9391643; PIDN:BAA02742.1; PID:9391644
A; Experimental source: stage 22-23 developing limb buds
A; Note: sequence extracted from NCBI backbone (NCBIN:134456, NCBIP:134457)
C; Superfamily: chicken chondroitin sulfate proteoglycan PG-M core protein; C-type 1e
F;166-243/Domain: link protein repeat homology <LNK1>
F;264-345/Domain: link protein repeat homology <LNK2>
F;3258-3289/Domain: EGF homology <EGF1>
F;3296-3327/Domain: EGF homology <EGF>
F;3334-3454/Domain: C-type lectin homology <LCH>
F;3461-3517/Domain: complement factor H repeat homology <FHD>
Query Match 15.3%; Score 171; DB 2; Length 3562;
Best Local Similarity 28.5%; Pred. No. 2.5e-06;
Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;
Db 3342 QGQCYK--YF--AHRRTWDAEERECLQAHLTSILSHEEQMFVNRV----GHDYQ 3389
Qy 76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWFVDEP----SCGSEVCVVMYHOPSAF 131
Db 3390 WIGL----NDKMFEEHDFRWTGDSLQYENWRPNQPDSSFFSAGEDCVCVIIWHENG-- 3439
Qy 132 AGIGGPMQWNDDRCNMKNNFICKYS---DEKPAVPSREAEGE 172
Db 3440 -QWNDDVPCNYHLLTYTCKKGTVACGQPPVVENAKTFFGK 3475
RESULT 10
T42710
mannose receptor, macrophage - mouse
N; Alternate names: lambda lectin; phospholipase A2 receptor
C; Species: Mus musculus (house mouse)
C; Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 *text_change 09-Jun-2000

C; Accession: T42710
 R; Wu, K.; Yuan, J.; Lasky, L.A.
 J. Biol. Chem. 271, 21323-21330, 1996
 A; Title: Characterization of a novel member of the macrophage mannose receptor type C 1e
 A; Reference number: Z22235; MUID: 96355501; PMID: 8702911
 A; Accession: T42710
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: mRNA
 A; Residues: 1-1479 <WUR>
 A; Cross-references: EMBL:U56734; NID:91336073; PID:q1336074; PIDN: AAC52729_1
 C; Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II repeat
 C; Keywords: membrane protein; receptor
 F; 186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 15.38; Score 170.5; DB 2; Length 1479;
 Best Local Similarity 31.48; Pred. No. 9.8e-07;
 Matches 48; Conservative 22; Mismatches 48; Indels 35; Gaps 7;

Qy 7 SGQPVCRGGTQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIEKF 66
 Db 384 SWQPF----QGHCYRL----QAEKRSWQESKRACLRGGDDLLSIHSMAELEPITKQIK 433

Qy 67 NLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSIISQFRNWWVDEPS---CGSEVCVV 123
 Db 434 QEVE--ELWIGL----NDLKLQMNFNEWSDGSLVSFTWHWPFEPPNNFRDSLEDCVT 482

Qy 124 MYHQPSAPAGIGGGPYMFQWNDDRCNMKNNFICKYS---DEKPAPVPSREAEGEETE 175
 Db 483 IW----GPEG----RWNDSPCNCQSLPSICK 504

RESULT 11

A39808 proteoglycan core protein, cartilage - bovine (fragments)
 N; Alternate names: aggrecan; aggregating cartilage proteoglycan
 C; Species: Bos primigenius taurus (cattle)
 C; Date: 20-Mar-1992 #sequence_revision 23-Mar-1995 #text_change 13-Aug-1999

C; Accession: A34234; A27752; A39808; A27751; E29164; B27751; D27751; F27
 R; Antonsson, P.; Heinegard, D.; Oldberg, A.
 J. Biol. Chem. 264, 16170-16173, 1989
 A; Title: The keratan sulfate-enriched region of bovine cartilage proteoglycan consists of
 A; Reference number: A34234; MUID: 89380219; PMID: 2528543
 A; Accession: A34234
 A; Status: preliminary; deduced from a
 A; Molecule type: mRNA
 A; Residues: 128-621 <ANT>
 A; Cross-references: GB:J05028
 R; Oldberg, A.; Antonsson, P.; Heinegard, D.
 Biochem. J. 243, 255-259, 1987
 A; Title: The partial amino acid sequence of bovine cartilage proteoglycan, deduced from a
 A; Reference number: A27752; MUID: 87270630; PMID: 3111460
 A; Accession: A27752
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 622-1340 <OLD>
 R; Sandy, J.D.; Boynton, R.E.; Flannery, C.R.
 J. Biol. Chem. 266, 8198-8205, 1991
 A; Title: Analysis of the catabolism of aggrecan in cartilage explants by quantitation of
 A; Reference number: A39808; MUID: 91217051; PMID: 2022637
 A; Accession: A39808
 A; Status: preliminary; protein
 A; Residues: 1-28; 59-82; 131-137, 'QSET', 142-149; 196-207; 226-249; 1137-1143; 1252-1267; 1274-1
 R; Perin, J.P.; Bonnet, F.; Jolles, P.
 FEBS Lett. 206, 73-77, 1986
 A; Title: Structural relationship between link proteins and proteoglycan monomers.
 A; Reference number: A27751; MUID: 87005253; PMID: 3530809

RESULT 12

T42630 aggrecan - bovine

C; Species: Bos primigenius taurus (cattle)
 C; Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 04-Mar-2000
 C; Accession: T42630
 R; Hering, T.M.; Kollar, J.; Huijrh, T.D.
 submitted to the EMBL Data Library, September 1996
 A; Description: Complete coding sequence of bovine aggrecan: comparative structural
 A; Reference number: Z22182
 A; Accession: T42630
 A; Status: preliminary; translated from GB/EMBL/DDBJ
 A; Molecule type: mRNA
 A; Residues: 1-2327 <HER>
 A; Cross-references: EMBL:U76615; NID:91730259; PID:91730260; PIDN: AAB38524_1

C; Superfamily: articular chondrocytes
 C; Type: C-type lectin homology; complement factor H repeat homolog
 C; Keywords: chondroitin sulfate proteoglycan; extracellular matrix; glyc

Query Match

14.28; Score 158.5; DB 2; Length 2327;

Best Local Similarity 26.8%; Pred. No. 1.9e-05;
 Matches 48; Conservative 29; Mismatches 59; Indels 43; Gaps 10;

Qy 9 QPVCRGG---TQRPCYKVYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIEKF 65
 Db 1127 QKLCEEGWTKFQGHCYR--HFPD---RATWVDAESQRKQQSHLSSIVTPEQ---EFV 1177

Qy 66 ENLLPSDGF-WIGLRRREEKQSNSTACQDLYAWTDGSIISQFRNWWVDEP---SCGSEV 120
 Db 1178 NN---NAQDYQWIGL----NDKTIEGDFRWSDGHSLOFENWRPNQDPDNFFATGEDC 1226

Qy 121 CVVVYHQPSAPAGIGGGPYMFQWNDDRCNMKNNFICKYS---DEKPAPVPSREAEGEETE 175
 Db 1227 VVMIWHEKG----EWNNDVPCNYQLPFTCKKGTVACGEPVVEHARIFGOKKD 1274

RESULT 13

A39086

aggrecan precursor, cartilage long splice form [validated] - human
 N; Alternate names: chondroitin sulfate Proteoglycan 1; large aggregating proteoglyc.
 N; Contains: aggrecan cartilage short splice form
 C; Species: Homo sapiens (man)
 C; Date: 10-Sep-1999 #sequence_revision 01-Dec-2000 #text_change 08-Dec-2000

C; Accession: A39086; S50206; A43919; S46659; S66389; S66846; S62786; A34226; B43919; C43
 R; Doege, K.J.; Sasaki, M.; Kimura, T.; Yamada, Y.
 J. Biol. Chem. 266, 894-902, 1991
 A; Title: Complete coding sequence and deduced primary structure of the human cartilage 1
 A; Reference number: A39086; MUID:91093289; PMID:1985970
 A; Accession: A39086
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-2162, 2201-2329, 'A', 2392-2415 <DOE>
 A; Cross-references: GB:MM5172; NID:9178258; PIDN:AAA62824.1; PID:9178259
 R; Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
 Biochim. Biophys. Acta 1219, 613-622, 1994
 A; Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
 A; Reference number: S50206; MUID:95035091; PMID:7524681
 A; Accession: S50206
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 350-497 <GLU>
 A; Cross-references: EMBL:X80278; NID:9516295
 A; Note: this translation is not annotated in GenBank entry HSAGGREC, release 113.0
 R; Sandy, J.D.; Flannery, C.R.; Neame, P.J.; Lohmander, L.S.
 J. Clin. Invest. 89, 1512-1516, 1992
 A; Title: The structure of aggrecan fragments in human synovial fluid. Evidence for the 1
 domain.
 A; Reference number: A43919; MUID:92235266; PMID:1569188
 A; Accession: A43919
 A; Molecule type: protein
 A; Residues: 361-370, 'X', 372-373; 393-399, 'X', 401-407, 'X', 409 <SAN>
 A; Cross-references: PIDN:ABB2079.1; PID:9248844; PIDN:AAB22077.1; PIDN:AAE
 A; Experimental source: synovial fluid
 A; Note: sequences modified after extraction from NCBI backbone
 R; Barry, F.P.; Neame, P.J.; Sasse, J.; Pearson, D.
 Matrix Biol. 14, 323-328, 1994
 A; Title: Length variation in the keratan sulfate domain of mammalian aggrecan.
 A; Reference number: 146998; MUID:95128522; PMID:7827755
 A; Accession: S46659
 A; Molecule type: DNA
 A; Residues: 764-765, 'A', 767-846, 'V', 848-862, 'X', 864 <BAR>
 A; Cross-references: EMBL:S74659; NID:9807127; PIDN:AAC60643.1; PID:9807128
 A; Note: the authors translated the codon GAA for residue 803 as Ala and CYT for residue
 R; Ilic, M.Z.; Mok, M.T.; Williamson, O.D.; Campbell, M.A.; Hughes, C.E.; Handley, C.J.
 Arch. Biochem. Biophys. 322, 22-30, 1995
 A; Title: Catabolism of aggrecan by explant cultures of human articular cartilage in the
 A; Reference number: S66389; MUID:96004775; PMID:7574678
 A; Accession: S66389
 A; Status: preliminary
 A; Molecule type: protein
 A; Residues: 17-23; 24, 'X', 26-27; 393-401; 402-403 <IL1>
 R; Fosang, A.J.; Last, K.; Knaeuper, V.; Murphy, G.; Neame, P.J.
 FEBS Lett. 380, 17-20, 1996
 A; Title: Age-related changes in the content of the C-terminal region of aggrecan in huma
 A; Reference number: S62786; MUID:96190740; PMID:8611178
 A; Status: preliminary; nucleic acid sequence not shown; translation not shown
 A; Molecule type: mRNA
 A; Residues: 1778-1927, 'A', 1929-1963, 'V', 1965-2162, 2201-2415 <DDU>
 A; Cross-references: EMBL:X17406; NID:930248; PIDN:CAA35463.1; PID:930249
 A; Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990
 R; Baldwin, C.T.; Reginato, A.M.; Prockop, D.J.
 J. Biol. Chem. 264, 15747-15750, 1989
 A; Title: A new epidermal growth factor-like domain in the human core protein for the lar
 A; Reference number: A34226; MUID:89380154; PMID:2789216
 A; Accession: A34226
 A; Molecule type: mRNA
 A; Residues: 1936-1963, 'V', 1965-2069, 'A', 2071-2415 <BAL>

A; Cross-references: GB:J05062; NID:9181167; PIDN:AAA35726.1; PID:9181168
 C; Genetics:
 A; Gene: GDB:AGC1; CSPG1; CSPGCP; MSK16
 A; Cross-references: GDB:127479; OMIM:155760
 A; Map position: 15q26-15q26
 C; Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology
 C; Keywords: alternative splicing; cartilage; chondroitin sulfate proteoglycan; extra
 F; 1-19/Domain: signal sequence *status predicted <SIG>
 F; 20-2415/Product: aggrecan cartilage long splice form *status Predicted <MATL>
 F; 20-2162, 2201-2415/Product: aggrecan cartilage short splice form *status Predicted <IMM>
 F; 44-135/Domain: immunoglobulin homology <IMM>
 F; 170-247/Domain: link protein repeat homology <LNK1>
 F; 268-349/Domain: link protein repeat homology <LNK2>
 F; 495-572/Domain: link protein repeat homology <LNK3>
 F; 593-673/Domain: keratan sulfate attachment *status Predicted <CS1>
 F; 677-861/Domain: chondroitin sulfate attachment *status Predicted <CS2>
 F; 1511-2162/Domain: chondroitin sulfate attachment *status Predicted <CS3>
 F; 2168-2198/Domain: EGF homology <EGF>
 F; 2205-2325/Domain: C-type lectin homology <LCH>
 F; 2332-2388/Domain: complement factor H repeat homology <FH>
 F; 126, 239, 333, 387, 434, 602, 657, 737, 1898/Binding site: carbohydrate (Asn) (covalent) *status Predicted <FHD>
 F; 371, 376/Binding site: keratan sulfate (Thr) (covalent) *status Predicted <FHD>
 Query Match 14.2%; Score 158.5; DB 1; Length 2415;
 Best Local Similarity 29.1%; Pred. No. 2e-05;
 Matches 52; Conservative 22; Mismatches 62; Indels 43; Gaps 11;
 Qy 9 QPVCRGG---TQRPCYKVYFHDTSRRLNFFEEAKEACRGQQLVSSIESEDEQKLIEKFIL 65
 Db 2202 QEVCEEGWNKQYQGHCYR---HFPD---RETWVDAERRCREEQQSHLSSIVTPEQ---EFV 2252
 Qy 66 ENLLPSPGDGDF-WIGLRRREEQKQNSNSTACQDLYAWTDGSISQFRNWYDEPS---CGSEVC 121
 Db 2253 NN---NAQDQWIGL---NDRTIEGDFRWSDGHMQFENWRPNQPDNFFAAGEDC 2301
 Qy 122 VVM-YHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAEGETE 175
 Db 2302 VVMIWHEKG-----EWNDVPCNYHLPFTCKKGTVACGGEPPVVEHARTFGQKKD 2349
 RESULT 14
 B42755
 E-selectin precursor - mouse
 N; Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)
 C; Species: Mus musculus (house mouse)
 C; Date: 04-Mar-1993 #sequence_revision 31-Dec-2000 #text_change 02-Aug-2002
 C; Accession: S23174; B42755
 R; Becker-Andre, M.; van Huitjdsdijnen, R.H.; Losberger, C.; Whelan, J.; Delamarre, J.
 Eur. J. Biochem. 206, 401-411, 1992
 A; Title: Murine endothelial leukocyte-adhesion molecule 1 is a close structural and
 A; Reference number: S23174; MUID:92283265; PMID:1375914
 A; Accession: S23174
 A; Status: preliminary
 A; Molecule type: DNA
 A; Residues: 1-612 <BEC>
 A; Cross-references: GB:MB00778; NID:9193014; PIDN:AAA37547.1; PID:9193015
 R; Weller, A.; Isenmann, S.; Vestweber, D.
 J. Biol. Chem. 267, 15176-15183, 1992
 A; Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin
 A; Reference number: A42755; MUID:92340571; PMID:1378846
 A; Accession: B42755
 A; Status: nucleic acid sequence not shown
 A; Molecule type: mRNA
 A; Residues: 'MKATAGV', 1-389, 391-612 <WEL>
 A; Cross-references: GB:MB87862; NID:9193107
 A; Experimental source: endothelial cells
 A; Note: sequence extracted from NCBI backbone (NCBIP:109470)
 A; Note: the sequence in GenBank entry MUSESELEC, release 117.0, (PIDN:AAA37577.1; PI
 A; Note: it is uncertain whether the initiator is Met-1 or the AUG codon preceding the
 C; Superfamily: C-type lectin homology; complement factor H repeat homolo
 C; Keywords: glycoprotein; transmembrane protein

F;1-21/Domain: signal sequence #status predicted <SIG>

F;22-612/Product: C-type lectin homology <LCH>

F;22-612/Product: P-selectin #status predicted <MAT>

F;143-174/Domain: EGF homology <EGF>

F;243-308/Domain: complement factor H repeat homology <FH1>

F;243-308/Domain: complement factor H repeat homology <FH2>

F;305-363/Domain: complement factor H repeat homology <FH3>

F;368-426/Domain: complement factor H repeat homology <FH4>

F;431-489/Domain: complement factor H repeat homology <FH5>

F;494-548/Domain: complement factor H repeat homology <FH6>

F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.8%; Score 153.5; DB 2; Length 2124;
 Best Local Similarity 26.3%; Pred. No. 4.8e-05;
 Matches 47; Conservative 27; Mismatches 62; Indels 43; Gaps 9;

Qy 9 QPYCRGG--TQRPCYKVIVFHDTSRRLNFEAKEACRRDGGOLVSISSDEQKLIEKF1 65
 Db 1911 QEQCCEEGWTKFQGHCYR--HFPD---RETWDAERRCREQQSHLSSIVTPEEQFVNKNA 1965

Qy 66 ENLLPSDGF-WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP---SCGSEV 120
 Db 1966 Q-----NDRTIEGFRWSGDGHSLQFEKWRPNQDPDNFFATGEDC 2010

Query Match 13.9%; Score 154.5; DB 2; Length 612;
 Best Local Similarity 27.9%; Pred. No. 9.2e-06;
 Matches 38; Conservative 28; Mismatches 45; Indels 25; Gaps 5;

Qy 25 YFHDTSRRLNFEAKEACRRDGGOLVSISSDEQKLIEKF1ENLLPSDGDFWIGLRRREE 84
 Db 23 YNNASSELMTYDEASAYCQRDYTHLVAQNKEE--INYLNNSNLKHSPSYYWIGRK--- 76

Qy 85 KQSNSTACQDLYAWTDGS--ISQFRNWYVDEPS--CGSEVCVVMYHQPSAPAGIGGPM 139
 Db 77 -----VNNVWIWVGTGKPLTEEAQNWAPEPNPKQRNEDCVEIYIQRTRDSGM--- 124

Qy 140 FQWNNDRCNMKNNFIC 155
 Db 125 -WNDERCNKKKLALC 138

RESULT 15

A28452

Proteoglycan core protein precursor, cartilage - rat

N;Alternate names: aggrecan

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000

C;Accession: A92623; A23835; A28095; A28452

R;Doege, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.

J. Biol. Chem. 262, 17757-17767, 1987

A;Title: Complete primary structure of the rat cartilage proteoglycan core protein deduced

A;Reference number: A92623; MUID:88087070; PMID:3693370

A;Accession: A92623

A;Molecule type: mRNA

A;Residues: 1-2124 <DOE>

R;Doege, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.

J. Biol. Chem. 263, 10040a, 1988

A;Reference number: A30069

A;Contents: annotation; revision to residue 698

R;Doege, K.; Hernandez, P.; Hassell, J.R.; Sasaki, M.; Yamada, Y.

J. Biol. Chem. 261, 8108-8111, 1986

A;Title: Partial cDNA sequence encoding a globular domain at the C terminus of the rat c

A;Reference number: A23835; MUID:86250698; PMID:2424893

A;Accession: A23835

A;Molecule type: protein

A;Residues: 1856-2124 <DOE>

A;Cross-references: GB: M13518; NID: g206104; PID: AAA41836.1; PMID: g206105

R;Neame, P.J.; Christner, J.E.; Baker, J.R.

J. Biol. Chem. 262, 17768-17778, 1987

A;Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-termin

A;Reference number: A28453; MUID:88087071; PMID:3693371

A;Accession: A28453

A;Molecule type: protein

A;Residues: 20-37, 'W', 39-60, 'E', 62-64, 'X', 66-69, 70-83; 84, 89-148, 'L', 150-238, 'S', 240, 'A'

C;Superfamily: aggrecan; C-type lectin homology; complement factor H repeat homology; EG

C;Keywords: glycoprotein

F;1-19/Domain: signal sequence #status predicted <SIG>

F;20-2124/Product: proteoglycan core protein #status predicted <MAT>

F;44-135/Domain: immunoglobulin homology <IMM>

F;170-247/Domain: link protein repeat homology <LNK1>

F;268-349/Domain: link protein repeat homology <LNK2>

F;504-581/Domain: link protein repeat homology <LNK3>

F;602-683/Domain: link protein repeat homology <LNK4>

F;1914-2034/Domain: C-type lectin homology <LCH>

F;2041-2097/Domain: complement factor H repeat homology <FHDC>

F;126, 239, 333, 387, 611, 667, 1842/Binding site: carbohydrate (Asn) (covalent) #status

F;22-612/Product: C-type lectin homology <LCH>

F;143-174/Domain: EGF homology <EGF>

F;243-308/Domain: complement factor H repeat homology <FH1>

F;305-363/Domain: complement factor H repeat homology <FH2>

F;368-426/Domain: complement factor H repeat homology <FH3>

F;431-489/Domain: complement factor H repeat homology <FH4>

F;494-548/Domain: complement factor H repeat homology <FH5>

F;25,391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

Copyright (c) 1993 - 2003 Compugen Ltd.

run on: May 20 2003 17:00:15 : 84

on: May 20, 2003, 17:00:13 ; Search Time 13 seconds
perfect score: 1115
sequence: 1 ATGRLLSGQPVCRGQTQRPC..... EEDAKKTFKESREAA
n: (without alignments)
657.240 Million cell updates/s

scoring table: BLOSUM62

卷之三

תרכזים: סדרה של מסמכי מסע, מ-1920 ל-1930, מילון של מילים עבריות.

total number of hits satisfying chosen parameters:

אָמַרְתִּי לְמַתְּהָרָה בְּבָבָלָן: זֶה מִתְּהָרָה!

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

database : SwissProt_40:*

red. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	574	51.5	273	1 CHOD_MOUSE	Q9CXM0 mus musculu
2	561.5	50.4	273	1 CHOD_HUMAN	Q9H9P2 homo sapien
3	182	16.3	1456	1 MANR_HUMAN	P22897 homo sapien
4	177	15.9	1268	1 PGCN_MOUSE	P55066 mus musculu
5	174.5	15.7	3381	1 PGCV_BOVIN	P81282 bos taurus
6	174	15.6	1257	1 PGCN_RAT	P55067 rattus norv
7	174	15.6	2738	1 PGCV_RAT	Q9erb4 rattus norv
8	174	15.6	3358	1 PGCV_MOUSE	Q62059 mus musculu
9	174	15.6	3396	1 PGCV_HUMAN	P13611 homo sapien
10	171	15.3	3562	1 PGCV_CHICK	Q90953 gallus gall
11	165	14.8	643	1 CD93_RAT	Q9et61 rattus norv
12	158.5	14.2	2364	1 PGCA_BOVIN	P13608 bos taurus
13	158.5	14.2	2415	1 PGCA_HUMAN	P16112 homo sapien
14	155.5	13.9	2333	1 PGCA_CANFA	Q28343 canis famil
15	154.5	13.9	612	1 LEM2_MOUSE	Q00690 mus musculu
16	153.5	13.8	644	1 CD93_MOUSE	Q89103 mus musculu
17	153.5	13.8	2124	1 PGCA_RAT	P07897 rattus norv
18	152	13.6	912	1 PGCB_BOVIN	Q28062 bos taurus
19	151.5	13.6	652	1 CD93_HUMAN	Q9nPY3 homo sapien
20	151	13.5	197	1 CLF1_HUMAN	Q75596 homo sapien
21	149	13.4	321	1 FCE2_HUMAN	P06734 homo sapien
22	148.5	13.3	2132	1 PGCA_MOUSE	Q61282 mus musculu
23	147	13.2	883	1 PGCB_MOUSE	Q61361 mus musculu
24	146	13.1	2109	1 PGCA_CHICK	P07898 gallus gall
25	145	13.0	158	1 LECG_TRIST	Q9YGP1 trimeresuru
26	145	13.0	883	1 PGCB_RAT	P55068 rattus norv
27	144.5	13.0	173	1 LEC2_MEGRO	P17346 megabalanus
28	144.5	13.0	372	1 LEM1_RAT	P30836 rattus norv
29	143.5	12.9	372	1 LEM1_MOUSE	P18337 mus musculu
30	141.5	12.7	331	1 FCE2_MOUSE	P20693 mus musculu
31	141.5	12.7	549	1 LEM2_RAT	P98105 rattus norv
32	140.5	12.6	162	1 LEC3_MEGRO	P07439 megabalanus
33			370	1 LEM1_BOVIN	P98131 bovin

34	139	12.5	PSPA_HUMAN	P07714	homo sapien
35	139	12.5	LECA_SARPE	P05047	sarcophaga
36	138.5	12.4	IXA_TRIFL	P23806	trimeresurus
37	138.5	12.4	TETN_MOUSE	P43025	mus musculu
38	137.5	12.3	LEM1_MACMU	Q95198	macaca mula
39	137.5	12.3	LEM1_PAPHA	Q28768	papio hamad
40	136	12.2	LITH_BOVIN	P23132	bos taurus
41	135	12.1	LIT1_MOUSE	P43137	mus musculu
42	134.5	12.1	ACAL_ANSAN	P83300	anser anser
43	134.5	12.1	LEM1_PONPY	Q95235	pongo pygma
44	134.5	12.1	LEM2_BOVIN	P98107	bos taurus
45	134	12.0	TETN_CARSP	P26258	carcharhinus

ALIGNMENT S

searched: 112892 seqs, 41476328 residues
total number of hits satisfying chosen parameters: 112892

minimum	DB	seq	length:
maximum	DB	seq	length:

post-processing: Minimum Match 0%

Maximizing water use
listing first 45 summaries

database : swissProt 10.1

卷之三

core greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases
 [2] SEQUENCE FROM N.A.
 STRAIN-C57BL/6J; TISSUE-Embryonic head;
 MEDLINE-21085660; PubMed-11217851;
 Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Ya-
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casav-
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochi
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Qu-
 Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., B.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonal-
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Garibo-
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momb-
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y.,
 Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Ko-
 Hayashizaki Y.;
 "Functional annotation of a full-length mouse cDNA collec-
 tion" Nature 409:685-690 (2001).
 -!- SUBCELLULAR LOCATION: Type I membrane protein (Poten-
 -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through
 CC between the Swiss Institute of Bioinformatics and the
 CC European Bioinformatics Institute. There are no re-
 CC use by non-profit institutions as long as its content
 CC is not modified and this statement is not removed. Usage by
 CC entities requires a license agreement (see <http://www.1sb.ch>)
 CC or send an email to license@1sb-sib.ch).

Hayashizaki Y.;
"Functional annotation of a full-length mouse cDNA collection. ";
Nature 409: 685-690 (2001).
-!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
-!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF311699; AAL50354.1; -.

DR	EMBL; AK014255; BAB29226.1; - .
DR	HSSP; P22897; 1EGG.
DR	MGD; MGI:1920461; 3110074E07Rik.
DR	InterPro; IPR001304; Lectin_C.
DR	Pfam; PF00059; lectin_c; 1.
DR	SMART; SM00034; CLECT; 1.
DR	PROSITE; PS00615; C_TYPELECTIN_1; FALSE_NEG.
DR	PROSITE; PS50041; C_TYPELECTIN_2; 1.
KW	Lectin; Transmembrane; Glycoprotein; Signal.
FT	SIGNAL 1 21 POTENTIAL.
FT	CHAIN 22 273 CHONDROLECTIN.
FT	DOMAIN 22 216 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 217 237 POTENTIAL.
FT	DOMAIN 238 273 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 35 179 C-TYPE LECTIN.
FT	CARBOHYD 86 86 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT 24 24 V -> W (IN REF. 2).
FT	CONFLICT 179 179 T -> K (IN REF. 2).
SQ	SEQUENCE 273 AA; 30303 MW; E052D933F244F4C7 CRC64;
Query Match	51.58; Score 574; DB 1; Length 273;
Best Local Similarity	57.48; Pred. No. 4e-44;
Matches 109; Conservative	26; Mismatches 41; Indels 14; Gaps 5;
Qy	4 RLLSGQPVCRGGTQRPCYKVIIYFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKLIKE 63
Db	23 RVVSGQQVKVCFADVKHPCYKMAKFHELSSRVSFQEARLACESEGGVLLSLENEAEQKLIES 82
Qy	64 FIENLLP----SDGDFWIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVDEPSCGS 118
Db	83 MLQNLTKPCTGDISDGDFWIGLLRSGDGT-SGACPDLYQWSDGSSSQFRNWYTDPEPSCGS 141
Qy	119 EVCVVWYHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYSDE-KPAVPSREAEGETELT 177
Db	142 EKCVVWYHQPTANPGLGGPYLYQWNNDDRCNMKHNHYICTYEPEIHPTEPA----EKPYLT 196
Qy	178 TPVLPETQE 187
Db	197 NQ--PEETHE 204 .
RESULT 2	
CHOD_HUMAN	STANDARD; PRT; 273 AA.
ID	Q9H9P2; Q9HCY3;
AC	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Chondrolectin precursor (Transmembrane protein MT75) (PREDI2
DE	protein).
GN	CHODL OR C21ORF68.
OS	<i>Homo sapiens</i> (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
OC	Mammalia; Eutherria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID	9606; [1]
RN	SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP	PubMed-12079284;
RX	Weng L., Smits P., Wauters J., Merregaert J.;
RA	"Molecular cloning and characterization of human chondrolectin, a novel type I transmembrane protein homologous to C-type lectins.";
RT	Genomics 80:62-70 (2002).
RL	[2]
RN	SEQUENCE FROM N.A.
RP	Medline-20289799; PubMed-10830953;
RX	Hattori M., Fujiiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W., Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E., Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,

RESULT 3			
	MANR_HUMAN	STANDARD;	PRT;
ID	P22897;		1456 AA.
AC			
DT	01-AUG-1991 (Rel. 19, Created)		
DT	01-AUG-1991 (Rel. 19, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Macrophage mannose receptor precursor (MMR) (CD206 antigen).		
GN	MRC1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX			
RN			
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RC	TISSUE=Placenta;		
RX	REVIEW; MEDLINE=90324192; PubMed=23373685;		
RA	Taylor M.E., Conary J.T., Lennartz M.R., Stahl P.D., Drickamer K.;		
RT	"Primary structure of the mannose receptor contains multiple motifs resembling carbohydrate-recognition domains.";		
RT	receptor (MRC1).";		
RL	J. Biol. Chem. 265:12156-12162(1990).		
RN			
RP	SEQUENCE FROM N.A.		
RX	REVIEW; MEDLINE=93052405; PubMed=12941118;		
RA	Kim S.J., Ruiz N., Bezouska K., Drickamer K.;		
RT	"Organization of the gene encoding the human macrophage mannose receptor (MRC1).";		
RT	Genomics 14:721-727(1992).		
RN			
RP	STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.		
RX	REVIEW; MEDLINE=92112893; PubMed=1730714;		
RA	Taylor M.E., Bezouska K., Drickamer K.;		
RT	"Contribution to ligand binding by multiple carbohydrate-recognition domains in the macrophage mannose receptor.";		
RT	J. Biol. Chem. 267:1719-1726(1992).		
RN			
RP	X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.		
RX	REVIEW; MEDLINE=20347275; PubMed=10779515;		
RA	Feinberg H., Park-Snyder S., Kollatkar A.R., Heise C.T., Taylor M.E.,		
RA	Weis W.I.;		
RT	"Structure of a C-type carbohydrate recognition domain from the macrophage mannose receptor.";		
RL	J. Biol. Chem. 275:21539-21548(2000).		
CC	- - FUNCTION: MEDIATES THE ENDOCYTOSIS OF GLYCOPROTEINS BY MACROPHAGES, IN SEVERAL RECOGNITION AND UPTAKE PROCESSES.		
CC	- - SUBCELLULAR LOCATION: Type I membrane protein.		
CC	- - MISCELLANEOUS: CRDS 1-3 HAVE AT MOST VERY WEAK AFFINITY FOR CARBOHYDRATE. CRD 4 SHOWS THE HIGHEST AFFINITY BINDING AND HAS MULTISPECIFICITY FOR A VARIETY OF MONOSACCHARIDES. AT LEAST 3 CRDS (4, 5, AND 7) ARE REQUIRED FOR HIGH AFFINITY BINDING AND ENDOCYTOSIS OF MULTIVALENT GLYCOCOCONJUGATES.		
CC	- - SIMILARITY: CONTAINS 8 C-TYPE LECTIN FAMILY DOMAINS.		
CC	- - SIMILARITY: CONTAINS 1 RICIN B-TYPE LECTIN DOMAIN.		
CC	- - DATABASE: NAME=PROW; NOTE=PROW 2:85-89 (2001);		
CC	WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1644341535_g.htm".		
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation, the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/anno		

Query Match 16.3%; Score 182; DB 1; Length 1456;
 Best Local Similarity 25.5%; Pred. No. 2.8e-08;
 Matches 52; Conservative 37; Mismatches 61; Indels 54; Gaps 8;

Qy 21 YKVIYFHDTSRRLNFEAKEACRRDGGOLVSISSEDEQKLIEKFIEENLLPSDGDFWIGLR 80
 Db 807 YKDYQYFSKEKETMDNARAFCKRNFGLVSIQSESEKKFLWKYY-NRNDAQSAFVIGLL 865

Qy 81 RREEKQSNSTACQDLYAYWTGDSISQFRNWWYDEPSGS--EVCVVMMYHQPSPAPAGIGGY 138
 Db 866 ISLDKK -----FAWDGSKVVDYVSWATGEPNFANEDENCVTMY---SNSGF---- 908

Qy 139 MFQWNDDRCNMKNNFICKYSDEK---PAVPSREAGEEETLTTPVLPETQE----- 187
 Db 909 ---WNIDINCYPNAFICQRHNSINATTVMP-----TMPSVSPSGCKEGWNFYSN 954

Qy 188 -----EDAKKTFKESREAA 202
 Db 955 KCFKIFFGFMEERKNTQEARAKACI 978

RESULT 4
 PGCN_MOUSE STANDARD; PRT; 1268 AA.

ID PGCN_MOUSE STANDARD; PRT; 1268 AA.
 AC P55066;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-Oct-2001 (Rel. 40, Last annotation update)
 DE Neurocan core protein precursor.
 GN CSPG3 OR NCAN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TAXID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RX MEDLINE=96039250; PubMed=7490074;
 RA Rauch U., Grimpe B., Kulbe G., Arnold-Ammer I., Beier D.,
 RA Faessler R.,
 RT "Structure and chromosomal localization of the mouse neurocan gene.";
 RL Genomics 28:405-410(1995).
 CC -1- FUNCTION: MAY MODULATE NEURONAL ADHESION AND NEURITE GROWTH DURING
 CC DEVELOPMENT BY BINDING TO NEURAL CELL ADHESION MOLECULES (NG-CAM
 CC AND N-CAM). CHONDROITIN SULFATE PROTEOGLYCAN; BINDS TO HYALURONIC
 CC ACID.
 CC -1- TISSUE SPECIFICITY: BRAIN.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

Query Match 15.9%; Score 177; DB 1; Length 1268;
 Best Local Similarity 31.5%; Pred. No. 6.6e-08;
 Matches 45; Conservative 17; Mismatches 49; Indels 32; Gaps 6;

Qy 17 QRPCYKVYYFDTSRRLNFEAKEACRRDGGOLVSISSEDEQKLIEKFIEENLLPSDGDFW 76
 Db 1048 QGHCYR--YF--AHRRAWEDAERDCCRRAAGHLTSVHSEEHKFINSF----GHENSW 1096

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWYDEPS---CGSEVCVVMYHQPSAPAG 133
 Db 1097 IGLNDRTVERD-----FQWTNTGLQYENREKOPDNFFAGGEDCVCVMVAHESG--- 1145

DR HSSP; P00740; 1EDM.
 DR MGD; MGI:104694; CSP93.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR000742; EGF-2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR001304; Lectin_C.
 DR InterPro; IPR000538; Link.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00008; EGF; 2.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF00059; lectin_C; 1.
 DR Pfam; PF00084; sushi; 1.
 DR Pfam; PF00193; Xlink; 2.
 DR PRODom; PD000918; Link; 2.
 DR SMART; SM00032; CCP; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00179; EGF_Ca; 1.
 DR SMART; SM00001; EGF-like; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00445; Link; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS0022; EGF-1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS01186; EGF-2; 1.
 DR PROSITE; PS01187; EGF_Ca; 1.
 DR PROSITE; PS01241; Link; 2.
 DR PROSITE; PS01242; Link; 2.
 DR PROSITE; PS01243; Link; 2.
 DR PROSITE; PS01244; Link; 2.
 DR PROSITE; PS01245; Link; 2.
 DR PROSITE; PS01246; Link; 2.
 DR PROSITE; PS01247; Link; 2.
 DR PROSITE; PS01248; Link; 2.
 DR PROSITE; PS01249; Link; 2.
 DR PROSITE; PS01250; Link; 2.
 DR PROSITE; PS01251; Link; 2.
 DR PROSITE; PS01252; Link; 2.
 DR PROSITE; PS01253; Link; 2.
 DR PROSITE; PS01254; Link; 2.
 DR PROSITE; PS01255; Link; 2.
 DR PROSITE; PS01256; Link; 2.
 DR PROSITE; PS01257; Link; 2.
 DR PROSITE; PS01258; Link; 2.
 DR PROSITE; PS01259; Link; 2.
 DR PROSITE; PS01260; Link; 2.
 DR PROSITE; PS01261; Link; 2.
 DR PROSITE; PS01262; Link; 2.
 DR PROSITE; PS01263; Link; 2.
 DR PROSITE; PS01264; Link; 2.
 DR PROSITE; PS01265; Link; 2.
 DR PROSITE; PS01266; Link; 2.
 DR PROSITE; PS01267; Link; 2.
 DR PROSITE; PS01268; Link; 2.
 DR PROSITE; PS01269; Link; 2.
 DR PROSITE; PS01270; Link; 2.
 DR PROSITE; PS01271; Link; 2.
 DR PROSITE; PS01272; Link; 2.
 DR PROSITE; PS01273; Link; 2.
 DR PROSITE; PS01274; Link; 2.
 DR PROSITE; PS01275; Link; 2.
 DR PROSITE; PS01276; Link; 2.
 DR PROSITE; PS01277; Link; 2.
 DR PROSITE; PS01278; Link; 2.
 DR PROSITE; PS01279; Link; 2.
 DR PROSITE; PS01280; Link; 2.
 DR PROSITE; PS01281; Link; 2.
 DR PROSITE; PS01282; Link; 2.
 DR PROSITE; PS01283; Link; 2.
 DR PROSITE; PS01284; Link; 2.
 DR PROSITE; PS01285; Link; 2.
 DR PROSITE; PS01286; Link; 2.
 DR PROSITE; PS01287; Link; 2.
 DR PROSITE; PS01288; Link; 2.
 DR PROSITE; PS01289; Link; 2.
 DR PROSITE; PS01290; Link; 2.
 DR PROSITE; PS01291; Link; 2.
 DR PROSITE; PS01292; Link; 2.
 DR PROSITE; PS01293; Link; 2.
 DR PROSITE; PS01294; Link; 2.
 DR PROSITE; PS01295; Link; 2.
 DR PROSITE; PS01296; Link; 2.
 DR PROSITE; PS01297; Link; 2.
 DR PROSITE; PS01298; Link; 2.
 DR PROSITE; PS01299; Link; 2.
 DR PROSITE; PS01300; Link; 2.
 DR PROSITE; PS01301; Link; 2.
 DR PROSITE; PS01302; Link; 2.
 DR PROSITE; PS01303; Link; 2.
 DR PROSITE; PS01304; Link; 2.
 DR PROSITE; PS01305; Link; 2.
 DR PROSITE; PS01306; Link; 2.
 DR PROSITE; PS01307; Link; 2.
 DR PROSITE; PS01308; Link; 2.
 DR PROSITE; PS01309; Link; 2.
 DR PROSITE; PS01310; Link; 2.
 DR PROSITE; PS01311; Link; 2.
 DR PROSITE; PS01312; Link; 2.
 DR PROSITE; PS01313; Link; 2.
 DR PROSITE; PS01314; Link; 2.
 DR PROSITE; PS01315; Link; 2.
 DR PROSITE; PS01316; Link; 2.
 DR PROSITE; PS01317; Link; 2.
 DR PROSITE; PS01318; Link; 2.
 DR PROSITE; PS01319; Link; 2.
 DR PROSITE; PS01320; Link; 2.
 DR PROSITE; PS01321; Link; 2.
 DR PROSITE; PS01322; Link; 2.
 DR PROSITE; PS01323; Link; 2.
 DR PROSITE; PS01324; Link; 2.
 DR PROSITE; PS01325; Link; 2.
 DR PROSITE; PS01326; Link; 2.
 DR PROSITE; PS01327; Link; 2.
 DR PROSITE; PS01328; Link; 2.
 DR PROSITE; PS01329; Link; 2.
 DR PROSITE; PS01330; Link; 2.
 DR PROSITE; PS01331; Link; 2.
 DR PROSITE; PS01332; Link; 2.
 DR PROSITE; PS01333; Link; 2.
 DR PROSITE; PS01334; Link; 2.
 DR PROSITE; PS01335; Link; 2.
 DR PROSITE; PS01336; Link; 2.
 DR PROSITE; PS01337; Link; 2.
 DR PROSITE; PS01338; Link; 2.
 DR PROSITE; PS01339; Link; 2.
 DR PROSITE; PS01340; Link; 2.
 DR PROSITE; PS01341; Link; 2.
 DR PROSITE; PS01342; Link; 2.
 DR PROSITE; PS01343; Link; 2.
 DR PROSITE; PS01344; Link; 2.
 DR PROSITE; PS01345; Link; 2.
 DR PROSITE; PS01346; Link; 2.
 DR PROSITE; PS01347; Link; 2.
 DR PROSITE; PS01348; Link; 2.
 DR PROSITE; PS01349; Link; 2.
 DR PROSITE; PS01350; Link; 2.
 DR PROSITE; PS01351; Link; 2.
 DR PROSITE; PS01352; Link; 2.
 DR PROSITE; PS01353; Link; 2.
 DR PROSITE; PS01354; Link; 2.
 DR PROSITE; PS01355; Link; 2.
 DR PROSITE; PS01356; Link; 2.
 DR PROSITE; PS01357; Link; 2.
 DR PROSITE; PS01358; Link; 2.
 DR PROSITE; PS01359; Link; 2.
 DR PROSITE; PS01360; Link; 2.
 DR PROSITE; PS01361; Link; 2.
 DR PROSITE; PS01362; Link; 2.
 DR PROSITE; PS01363; Link; 2.
 DR PROSITE; PS01364; Link; 2.
 DR PROSITE; PS01365; Link; 2.
 DR PROSITE; PS01366; Link; 2.
 DR PROSITE; PS01367; Link; 2.
 DR PROSITE; PS01368; Link; 2.
 DR PROSITE; PS01369; Link; 2.
 DR PROSITE; PS01370; Link; 2.
 DR PROSITE; PS01371; Link; 2.
 DR PROSITE; PS01372; Link; 2.
 DR PROSITE; PS01373; Link; 2.
 DR PROSITE; PS01374; Link; 2.
 DR PROSITE; PS01375; Link; 2.
 DR PROSITE; PS01376; Link; 2.
 DR PROSITE; PS01377; Link; 2.
 DR PROSITE; PS01378; Link; 2.
 DR PROSITE; PS01379; Link; 2.
 DR PROSITE; PS01380; Link; 2.
 DR PROSITE; PS01381; Link; 2.
 DR PROSITE; PS01382; Link; 2.
 DR PROSITE; PS01383; Link; 2.
 DR PROSITE; PS01384; Link; 2.
 DR PROSITE; PS01385; Link; 2.
 DR PROSITE; PS01386; Link; 2.
 DR PROSITE; PS01387; Link; 2.
 DR PROSITE; PS01388; Link; 2.
 DR PROSITE; PS01389; Link; 2.
 DR PROSITE; PS01390; Link; 2.
 DR PROSITE; PS01391; Link; 2.
 DR PROSITE; PS01392; Link; 2.
 DR PROSITE; PS01393; Link; 2.
 DR PROSITE; PS01394; Link; 2.
 DR PROSITE; PS01395; Link; 2.
 DR PROSITE; PS01396; Link; 2.
 DR PROSITE; PS01397; Link; 2.
 DR PROSITE; PS01398; Link; 2.
 DR PROSITE; PS01399; Link; 2.
 DR PROSITE; PS01400; Link; 2.
 DR PROSITE; PS01401; Link; 2.
 DR PROSITE; PS01402; Link; 2.
 DR PROSITE; PS01403; Link; 2.
 DR PROSITE; PS01404; Link; 2.
 DR PROSITE; PS01405; Link; 2.
 DR PROSITE; PS01406; Link; 2.
 DR PROSITE; PS01407; Link; 2.
 DR PROSITE; PS01408; Link; 2.
 DR PROSITE; PS01409; Link; 2.
 DR PROSITE; PS01410; Link; 2.
 DR PROSITE; PS01411; Link; 2.
 DR PROSITE; PS01412; Link; 2.
 DR PROSITE; PS01413; Link; 2.
 DR PROSITE; PS01414; Link; 2.
 DR PROSITE; PS01415; Link; 2.
 DR PROSITE; PS01416; Link; 2.
 DR PROSITE; PS01417; Link; 2.
 DR PROSITE; PS01418; Link; 2.
 DR PROSITE; PS01419; Link; 2.
 DR PROSITE; PS01420; Link; 2.
 DR PROSITE; PS01421; Link; 2.
 DR PROSITE; PS01422; Link; 2.
 DR PROSITE; PS01423; Link; 2.
 DR PROSITE; PS01424; Link; 2.
 DR PROSITE; PS01425; Link; 2.
 DR PROSITE; PS01426; Link; 2.
 DR PROSITE; PS01427; Link; 2.
 DR PROSITE; PS01428; Link; 2.
 DR PROSITE; PS01429; Link; 2.
 DR PROSITE; PS01430; Link; 2.
 DR PROSITE; PS01431; Link; 2.
 DR PROSITE; PS01432; Link; 2.
 DR PROSITE; PS01433; Link; 2.
 DR PROSITE; PS01434; Link; 2.
 DR PROSITE; PS01435; Link; 2.
 DR PROSITE; PS01436; Link; 2.
 DR PROSITE; PS01437; Link; 2.
 DR PROSITE; PS01438; Link; 2.
 DR PROSITE; PS01439; Link; 2.
 DR PROSITE; PS01440; Link; 2.
 DR PROSITE; PS01441; Link; 2.
 DR PROSITE; PS01442; Link; 2.
 DR PROSITE; PS01443; Link; 2.
 DR PROSITE; PS01444; Link; 2.
 DR PROSITE; PS01445; Link; 2.
 DR PROSITE; PS01446; Link; 2.
 DR PROSITE; PS01447; Link; 2.
 DR PROSITE; PS01448; Link; 2.
 DR PROSITE; PS01449; Link; 2.
 DR PROSITE; PS01450; Link; 2.
 DR PROSITE; PS01451; Link; 2.
 DR PROSITE; PS01452; Link; 2.
 DR PROSITE; PS01453; Link; 2.
 DR PROSITE; PS01454; Link; 2.
 DR PROSITE; PS01455; Link; 2.
 DR PROSITE; PS01456; Link; 2.
 DR PROSITE; PS01457; Link; 2.
 DR PROSITE; PS01458; Link; 2.
 DR PROSITE; PS01459; Link; 2.
 DR PROSITE; PS01460; Link; 2.
 DR PROSITE; PS01461; Link; 2.
 DR PROSITE; PS01462; Link; 2.
 DR PROSITE; PS01463; Link; 2.
 DR PROSITE; PS01464; Link; 2.
 DR PROSITE; PS01465; Link; 2.
 DR PROSITE; PS01466; Link; 2.
 DR PROSITE; PS01467; Link; 2.
 DR PROSITE; PS01468; Link; 2.
 DR PROSITE; PS01469; Link; 2.
 DR PROSITE; PS01470; Link; 2.
 DR PROSITE; PS01471; Link; 2.
 DR PROSITE; PS01472; Link; 2.
 DR PROSITE; PS01473; Link; 2.
 DR PROSITE; PS01474; Link; 2.
 DR PROSITE; PS01475; Link; 2.
 DR PROSITE; PS01476; Link; 2.
 DR PROSITE; PS01477; Link; 2.
 DR PROSITE; PS01478; Link; 2.
 DR PROSITE; PS01479; Link; 2.
 DR PROSITE; PS01480; Link; 2.
 DR PROSITE; PS01481; Link; 2.
 DR PROSITE; PS01482; Link; 2.
 DR PROSITE; PS01483; Link; 2.
 DR PROSITE; PS01484; Link; 2.
 DR PROSITE; PS01485; Link; 2.
 DR PROSITE; PS01486; Link; 2.
 DR PROSITE; PS01487; Link; 2.
 DR PROSITE; PS01488; Link; 2.
 DR PROSITE; PS01489; Link; 2.
 DR PROSITE; PS01490; Link; 2.
 DR PROSITE; PS01491; Link; 2.
 DR PROSITE; PS01492; Link; 2.
 DR PROSITE; PS01493; Link; 2.
 DR PROSITE; PS01494; Link; 2.
 DR PROSITE; PS01495; Link; 2.
 DR PROSITE; PS01496; Link; 2.
 DR PROSITE; PS01497; Link; 2.
 DR PROSITE; PS01498; Link; 2.
 DR PROSITE; PS01499; Link; 2.
 DR PROSITE; PS01500; Link; 2.
 DR PROSITE; PS01501; Link; 2.
 DR PROSITE; PS01502; Link; 2.
 DR PROSITE; PS01503; Link; 2.
 DR PROSITE; PS01504; Link; 2.
 DR PROSITE; PS01505; Link; 2.
 DR PROSITE; PS01506; Link; 2.
 DR PROSITE; PS01507; Link; 2.
 DR PROSITE; PS01508; Link; 2.
 DR PROSITE; PS01509; Link; 2.
 DR PROSITE; PS01510; Link; 2.
 DR PROSITE; PS01511; Link; 2.
 DR PROSITE; PS01512; Link; 2.
 DR PROSITE; PS01513; Link; 2.
 DR PROSITE; PS01514; Link; 2.
 DR PROSITE; PS01515; Link; 2.
 DR PROSITE; PS01516; Link; 2.
 DR PROSITE; PS01517; Link; 2.
 DR PROSITE; PS01518; Link; 2.
 DR PROSITE; PS01519; Link; 2.
 DR PROSITE; PS01520; Link; 2.
 DR PROSITE; PS01521; Link; 2.
 DR PROSITE; PS01522; Link; 2.
 DR PROSITE; PS01523; Link; 2.
 DR PROSITE; PS01524; Link; 2.
 DR PROSITE; PS01525; Link; 2.
 DR PROSITE; PS01526; Link; 2.
 DR PROSITE; PS01527; Link; 2.
 DR PROSITE; PS01528; Link; 2.
 DR PROSITE; PS01529; Link; 2.
 DR PROSITE; PS01530; Link; 2.
 DR PROSITE; PS01531; Link; 2.
 DR PROSITE; PS01532; Link; 2.
 DR PROSITE; PS01533; Link; 2.
 DR PROSITE; PS01534; Link; 2.
 DR PROSITE; PS01535; Link; 2.
 DR PROSITE; PS01536; Link; 2.
 DR PROSITE; PS01537; Link; 2.
 DR PROSITE; PS01538; Link; 2.
 DR PROSITE; PS01539; Link; 2.
 DR PROSITE; PS01540; Link; 2.
 DR PROSITE; PS01541; Link; 2.
 DR PROSITE; PS015

QY	134	IGGPYMFQWNDDRCNMKNNFICK	156
	:		:
Db	1146	-----RWNDVPCNYNLPYVCK	1161
RESULT 5			
	PGCV_BOVIN	STANDARD;	PRT; 3381 AA.
ID	P81282; 077609; 077610; 077611; 077612;		
AC	15-DEC-1998 (Rel. 37, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Versican core protein precursor (Large fibroblast proteoglycan (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glia		
DE	hyaluronate-binding protein) (GHAP).		
GN	CPSG2.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.		
OC	NCBI_TAXID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORMS V0; V1; V2 AND V3).		
RC	TISSUE-Forebrain;		
RX	MEDLINE-98288320; PubMed-9624174;		
RA	Schmalfieldt M., Dours-Zimmermann M.T., Winterhalter K.H., Zimmermann D.R.;		
RA	"versican V2 is a major extracellular matrix component of the mature bovine brain."		
RT	RT		
RL	J. Biol. Chem. 273:15758-15764 (1998).		
RN	[2]		
RP	SEQUENCE OF 21-53; 78-96; 226-250; 262-277; 295-306; 314-324; 329-331 AND 342-348.		
RC	TISSUE-Spinal cord;		
RX	MEDLINE-92062692; PubMed-1720020;		
RA	Perides G., Biviano F., Bignami A.;		
RA	"Interaction of a brain extracellular matrix protein with hyaluronic acid."		
RL	Biochim. Biophys. Acta 1075:248-258 (1991).		
CC	- - FUNCTION: May play a role in intercellular signaling and in connecting cells with the extracellular matrix. May take part in the regulation of cell motility, growth and differentiation. Binds hyaluronic acid.		
CC	- - SUBCELLULAR LOCATION: Secreted; extracellular matrix.		
CC	- - ALTERNATIVE PRODUCTS: At least 4 isoforms; v0 (shown here), v1, v2 and v3; are produced by alternative splicing.		
CC	- - TISSUE SPECIFICITY: Cerebral white matter. v0 and v1 are expressed in the central nervous system, and in a number of mesenchymal and epithelial tissues; the major isoform v2 is restricted to the central nervous system.		
CC	- - DEVELOPMENTAL STAGE: Disappears after the cartilage development (By similarity).		
CC	- - SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.		
CC	- - SIMILARITY: CONTAINS 2 LINK DOMAINS.		
CC	- - SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.		
CC	- - SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.		
CC	- - SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.		
CC	- - SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.		

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcements or send an email to license@isb-sib.ch).		
CC	EMBL; AF060456; AAC24358.1; -.		
CC	EMBL; AF060457; AAC24359.1; -.		
CC	EMBL; AF060458; AAC24360.1; -.		
CC	EMBL; AF060459; AAC24361.1; -.		
CC	HSSP; P01132; 1EPP.		
CC	InterPro; IPR000152; Asx_hydroxyL.		

PT CARBOHYD 2045 2045 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2074 2074 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2103 2103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2263 2263 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2290 2290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2356 2356 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2623 2623 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2641 2641 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 2919 2919 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3052 3052 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3354 3354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 3364 3364 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 349 349 P->R (IN ISOFORM V1 AND V3).
 FT VARSPLIC 350 1336 MISSING (IN ISOFORM V1).
 FT VARSPLIC 1337 3074 MISSING (IN ISOFORM V2).
 FT VARSPLIC 350 3074 MISSING (IN ISOFORM V3).
 FT CONFLICT 25 25 MISSING (IN REF. 2).
 FT CONFLICT 51 51 MISSING (IN REF. 2).
 FT CONFLICT 89 89 N->D (IN REF. 2).
 FT CONFLICT 96 96 Q->D (IN REF. 2).
 FT CONFLICT 346 346 C->R (IN REF. 2).
 SQ SEQUENCE 3381 AA; 369984 MW; F09716FA778D459 CRC64;

Query Match 15.7%; Score 174.5%; DB 1; Length 3381;
 Best Local Similarity 25.5%; Pred. No. 3.5e-07; Indels 75; Gaps 11;
 Matches 59; Conservative 26; Mismatches 71; Gaps 11;

QY 17 QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSISEDEQKLIEKFIELNLLPSDGDF- 75
 Db 31162 QGQCYK--YF---AHRRTWDAEERECLQGAHLTSILSHEEQMFVNVR----GHDYQ 3209

QY 76 WIGLRRREEKQNSNSTACQDLYAWTDGSISQFRNWYVDEP---SCGSEVCVVMYHOPSAAP 131
 Db 3210 WIGL--NDKMFHDFRWTGDSLQYENWRNPQDSFFSTGEDCVVIIWHENG-- 3259

QY 132 AGIGGPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAGE----- 172
 Db 3260 -----QWNDDVPCNYHLTYYTCKKGTVACGQPPVVENAKTFGKMKPRYEINSLIRYHC 3310

QY 173 -----ETELTT----PVL---PEETQEEDAKKTKESESREAAALN 203
 Db 3311 KDGFIQRLHPTIRLGNRAMPKITCLNPSAYQRTYSKKYFKNNSSSAKDN 3361

RESULT 6 PGCN_RAT
 ID PGCN_RAT STANDARD; PRT; 1257 AA.
 AC P55067;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurocan core protein precursor (245 kDa early postnatal core
 glycoprotein) (contains: 150 kDa adult core glycoprotein).
 GN CSPG3 OR NCAN.
 OS Rattus norvegicus (Rat).
 OC Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RN STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=92406907; PubMed=1326557;
 RA Rauch U., Karthikeyan L., Maurel P., Margolis R.U., Margolis R.K.,
 RT "Cloning and primary structure of neurocan, a developmentally
 regulated, aggregating chondroitin sulfate proteoglycan of brain."
 RL J. Biol. Chem. 267:19536-19547 (1992).
 RN [2] CHARACTERIZATION.
 RX MEDLINE=94230574; PubMed=7513709;
 RA Friedlander D.R., Milev P., Karthikeyan L., Margolis R.K.,
 RA Margolis R.U., Grumet M.;
 RT "The neuronal chondroitin sulfate proteoglycan binds to the
 neural cell adhesion molecules Ng-CAM/L1/NILE and N-CAM, and inhibits
 by SIMILARITY.
 RT DISULFID 181 252
 FT DISULFID 205 226

RT neuronal adhesion and neurite outgrowth.";
 J. Cell. Biol. 125:669-680 (1994).
 -I- FUNCTION: MAY MODULATE NEURONAL ADHESION AND NEURITE GROWTH DURING
 DEVELOPMENT BY BINDING TO NEURAL CELL ADHESION MOLECULES (NG-CAM
 AND N-CAM). CHONDROITIN SULFATE PROTEOGLYCAN; BINDS TO HYALURONIC
 ACID.
 -I- TISSUE SPECIFICITY: EARLY POSTNATAL AND ADULT BRAIN; NOT EXPRESSED
 IN KIDNEY, LUNG, LIVER AND MUSCLE.
 CC -I- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND
 O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).
 CC -I- PTM: TWO ISOFORMS WERE FOUND THAT PROBABLY ARISE BY PROTEOLYTIC
 DEGRADATION. THE LARGE ISOFORM IS PREDOMINANT IN EARLY POSTNATAL
 BRAIN, THE SMALL ISOFORM IS FOUND IN ADULT BRAIN.
 CC -I- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -I- SIMILARITY: CONTAINS 1 SCR DOMAIN.
 CC -I- SIMILARITY: BELONGS TO THE AGGREGCAN/VERSIKAN PROTEOGLYCAN FAMILY.
 CC -I- SIMILARITY: CONTAINS 2 EGFR-LIKE DOMAINS.
 CC -I- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -I- SIMILARITY: CONTAINS 1 C

FT DISULFID 279 354 BY SIMILARITY.
 FT DISULFID 303 324 BY SIMILARITY.
 FT DISULFID 953 964 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 975 984 BY SIMILARITY.
 FT DISULFID 1029 1040 BY SIMILARITY.
 FT DISULFID 1057 1149 BY SIMILARITY.
 FT DISULFID 1125 1141 BY SIMILARITY.
 FT DISULFID 1156 1199 BY SIMILARITY.
 FT DISULFID 1185 1212 BY SIMILARITY.
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 737 737 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 944 944 O-LINKED (XYL. .) (CHONDROITIN SULFATE).
 FT CARBOHYD 967 967 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1164 1164 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1257 AA; 135544 MW; 992B33DCFA19EE1B CRC64;

Query Match 15.6%; Score 174; DB 1; Length 1257;
 Best Local Similarity 30.8%; Pred. No. 1.2e-07;
 Matches 44; Conservative 18; Mismatches 49; Indels 32; Gaps 6;

Qy 17 QRPCYKVIFHDTSRRLNFEAAKEACRRDGGQVLVSIESEDEQKLIEKFIENLLPSDGDFW 76
 Db 1037 QGHCYR--YF--AHRRAWEDAERDCRRAGHLTSVHSPEEEHKFINSF----GHENSW 1085

Qy 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEPS---CGSEVCVVMYHQPSAPAG 133
 Db 1086 IGLNDRTVERD-----FQWTNDTGLQYENWREKQPDNEFFAGGEDCVVVMVAHENG--- 1134

Qy 134 IGGPYMFQWNDDRCNMKNNFICK 156
 Db 1135 -----RWNNDVPCNYNLPPYCK 1150

RESULT 7
 PGCV_RAT
 ID PGCV_RAT STANDARD; PRT; 2738 AA.
 AC Q9ERB4; 008592; Q88564; Q9R1K4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Versican core protein precursor (Large fibroblast proteoglycan)
 DE (Chondroitin sulfate proteoglycan core protein 2) (PG-M) (Glia)
 DE hyaluronate-binding protein) (GHAP) (Fragments).
 GN CPSG2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID-10116;
 RN [1] SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0), SEQUENCE FROM N.A.
 RP (ISOFORM V3), AND SEQUENCE OF 2657-2738 FROM N.A. (ISOFORM VINT).
 RC STRAIN=Wistar Kyoto;
 RX MEDLINE=99327053; PubMed=10397680;
 RA Lemire J.M., Braun K.R., Maurel P., Kaplan E.D., Schwartz S.M.,
 RA Wight T.N.;
 RT "Versican/PG-M Isoforms in vascular smooth muscle cells.";
 RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
 RN [2] SEQUENCE OF 349-2738 FROM N.A. (ISOFORM V0).
 RP STRAIN=Wistar Kyoto;
 RX MEDLINE=98308094; PubMed=9642104;
 RA Milev P., Maurel P., Chiba A., Mevissen M., Popp S., Yamaguchi Y.,
 RA Margolis R.K., Margolis R.U.;
 RT "Differential regulation of expression of hyaluronan-binding
 proteoglycans in developing brain: aggrecan, versican, neurocan, and
 brevican.";
 RL Biochem. Biophys. Res. Commun.. 247:207-212(1998).
 RN [3] SEQUENCE OF 2421-2463 FROM N.A. (ISOFORM V0).
 RC TISSUE=Kidney;
 RX MEDLINE=98094159; PubMed=9434070;

RA Pyke C., Kristensen P., Ostergaard P.B., Oturai P.S., Romer J.;
 RT "Proteoglycan expression in the normal rat kidney.";
 RL Nephron 77:461-470(1997).
 RN [4] SEQUENCE OF 2535-2738 FROM N.A.
 RP STRAIN-Sprague-Dawley; TISSUE=Lung;
 RC Bloomberg L.A., Chan W.-Y., Clerch L., Massaro D.;
 RT "Molecular cloning and characterization of two developmentally
 regulated genes in rat lung.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
 CC -!- FUNCTION: May play a role in intercellular signalling and in
 CC connecting cells with the extracellular matrix. May take part in
 CC the regulation of cell motility, growth and differentiation. Binds
 CC hyaluronic acid.
 CC -!- SUBCELLULAR LOCATION: Secreted; extracellular matrix. V0 (shown here), V3 and
 CC Vint; are produced by alternative splicing.
 CC -!- TISSUE SPECIFICITY: In kidney is expressed in the papillary area,
 CC but not in glomeruli.
 CC -!- DEVELOPMENTAL STAGE: Disappears after the cartilage development
 CC (By similarity).
 CC -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -!- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.
 CC
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR AF062402; AAC40166.1;
 DR EMBL; U75306; AAB51125.1;
 DR EMBL; AF084544; AAD48544.1;
 DR EMBL; AF072892; AAC26116.1;
 DR EMBL; AX007691; AAG16631.1;
 DR HSSP; P01132; 1EPG.
 DR InterPro; IPR000152;
 DR InterPro; IPR000561;
 DR InterPro; IPR000742;
 DR InterPro; IPR001881;
 DR InterPro; IPR003599;
 DR InterPro; IPR003600;
 DR InterPro; IPR001304;
 DR InterPro; IPR000538;
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR PRINTS; PR01265; LINKMODULE.
 DR SMART; SM00032; CCP; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 2.
 DR SMART; SM00179; EGF-CA; 2.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00410; IG_Like; 1.
 DR SMART; SM00445; LINK; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS00615; C_TYPELECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPELECTIN_2; 1.
 DR PROSITE; PS01186; EGF; 2.
 DR PROSITE; PS01187; EGF-CA; 1.
 DR PROSITE; PS01241; LINK; 2.
 KW Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;
 KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 KW Hyaluronic acid; Alternative splicing;
 FT SIGNAL 1 20
 FT CHAIN 21 2738
 FT NON_CONS 348 349
 FT DOMAIN 37 137

FT DOMAIN	167	244	LINK 1.	
FT DOMAIN	265	346	LINK 2.	
FT DOMAIN	<349	695	GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).	
FT DOMAIN	696	2431	GAG-BETA.	
FT DOMAIN	2431	2467	EGF-LIKE 1.	
FT DOMAIN	2469	2505	EGF-LIKE 2, CALCIUM-BINDING.	
FT DOMAIN	2518	2632	C-TYPE LECTIN.	
FT DISULFID	2637	2695	SUSHI.	
FT DISULFID	44	130	BY SIMILARITY.	
FT DISULFID	172	243	BY SIMILARITY.	
FT DISULFID	196	217	BY SIMILARITY.	
FT DISULFID	270	345	BY SIMILARITY.	
FT DISULFID	294	315	BY SIMILARITY.	
FT DISULFID	2435	2446	BY SIMILARITY.	
FT DISULFID	2440	2455	BY SIMILARITY.	
FT DISULFID	2457	2466	BY SIMILARITY.	
FT DISULFID	2473	2484	BY SIMILARITY.	
FT DISULFID	2478	2493	BY SIMILARITY.	
FT DISULFID	2495	2504	BY SIMILARITY.	
FT DISULFID	2511	2522	BY SIMILARITY.	
FT DISULFID	2539	2631	BY SIMILARITY.	
FT DISULFID	2607	2623	BY SIMILARITY.	
FT DISULFID	2638	2681	BY SIMILARITY.	
FT DISULFID	2667	2694	BY SIMILARITY.	
FT CARBOHYD	57	57	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	330	330	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	692	692	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	758	758	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	802	802	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	805	805	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1257	1257	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1435	1435	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1633	1633	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1660	1660	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1684	1684	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1738	1738	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	1848	1848	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	2004	2004	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	2409	2409	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	2711	2711	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT CARBOHYD	2721	2721	N-LINKED (GLCNAC. .) (POTENTIAL).	
FT VARSPLIC	349	2431	MISSING (IN ISOFORM V3).	
FT VARSPLIC	2697	2738	PSAYQRTYSKKYLNSSSSVQDNSTSINTSKHEHRWSRRQETR R -> RKWSFRKNGOPCFNKY (IN ISOFORM VINT).	
FT CONFLICT	2535	2539	AERECA -> NSARG (IN REF. 4).	
SQ SEQUENCE	2738	AA;	300004 MW; 12CA626D58BD8C6A CRC64;	
Query Match		15.6%	Score 174; DB 1; Length 2738;	
Best Local Similarity		28.5%	Pred. No. 3e-07;	
Matches 47; Conservative		23;	Mismatches 55; Indels 40; Gaps 8;	
Qy 17	QRPCYKVIYFHDTSRRLNFEEAKEACRRDGGOLYSIESEDEQKLIEKFENLLPSDGDF- 75			
Db 2519	QGQCYK--YF---AHRRTWDAARECRLQGAHLTSILSHEEQMFTVNRV-----GHDYQ 2566			
Qy 76	WIGLRRREEKOSNSNSTACQDLYAWTDGSIISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131			
Db 2567	WIGL-----NDKMFEHDFRWTGDSALQENWRNPQDSSFFSAGEDCVVIIWHENG-- 2616			
Qy 132	AGIGGPPMFQWNDDRCNMKNNFICKYS---DEKPAPVPSREAEGE 172			
Db 2617	-----QWNDDVPCNYHLTYTCKKGTVACGQQPVVNAKTFGK 2652			
RESULT 8				
PGCV_MOUSE	STANDARD;	PRT;	3358 AA.	
ID PGCV_MOUSE				
AC 062059; Q62058; Q9CUU0;				
DT 01-NOV-1997 (Rel. 35, Created)				
DT 01-NOV-1997 (Rel. 35, Last sequence update)				
DT 16-OCT-2001 (Rel. 40, Last annotation update)				
DE Versican core protein precursor (Large fibroblast proteoglycan)				
DR EMBL; D16263; BAA03796.1; -.				
EMBL; D28599; -; NOT_ANNOTATED_CDS.				

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; D16263; BAA03796.1; -.

DR	EMBL; D32040; BAA06802.1; -.	FT	CARBOHYD	1372	1372	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	EMBL; AK014525; BAB29411.1; -.	FT	CARBOHYD	1679	1679	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	HSSP; P01132; 1EPG.	FT	CARBOHYD	2054	2054	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	MGD; MGI:102889; CSPG2.	FT	CARBOHYD	2244	2244	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR000152; Asx_hydroxyl.	FT	CARBOHYD	2362	2362	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR000561; EGF-1 like.	FT	CARBOHYD	2627	2627	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR000742; EGF_2.	FT	CARBOHYD	3030	3030	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR001881; EGF_Ca.	FT	CARBOHYD	3332	3332	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR001438; EGF_II.	FT	CARBOHYD	3342	3342	N-LINKED (GLCNAC. .) (POTENTIAL).
DR	InterPro; IPR003599; Ig.	FT	VARSPLIC	348	348	P -> R (IN ISOFORM VI AND ISOFORM V3).
DR	InterPro; IPR003006; Ig_MHC.	FT	VARSPLIC	349	1308	MISSING (IN ISOFORM V1).
DR	InterPro; IPR001304; Lectin_C.	FT	VARSPLIC	1309	3052	MISSING (IN ISOFORM V2).
DR	InterPro; IPR000538; Link.	FT	VARSPLIC	349	3052	MISSING (IN ISOFORM V3).
DR	InterPro; IPR000436; Sushi_SCR_CCP.	FT	CONFFLICT	126	126	A -> G (IN REF. 3).
DR	Pfam; PF000008; EGF; 2.	FT	CONFFLICT	348	348	MISSING (IN REF. 3).
DR	Pfam; PF00047; Ig; 1.	FT	CONFFLICT	1658	1658	I -> T (IN REF. 3).
DR	Pfam; PF00059; Lectin_C; 1.	FT	CONFFLICT	1674	1680	TWWSNSNS -> QFGIOTA (IN REF. 3).
DR	Pfam; PF00084; Sushi; 1.	SQ	SEQUENCE	3358	AA; 366938 MW;	071B80026BC0762D CRC64;
DR	PRINTS; PRO0010; EGGFBLOOD.					
DR	PRODOM; PD000918; Link; 2.					
DR	SMART; SM00032; CCP; 1.					
DR	SMART; SM00034; CLECT; 1.					
DR	SMART; SM00179; EGF_Ca; 1.					
DR	SMART; SM00001; EGF-like; 1.					
DR	SMART; SM00409; Ig; 1.					
DR	SMART; SM00445; Link; 2.					
DR	PROSITE; PS000010; ASX_HYDROXYL; 1.					
DR	PROSITE; PS00022; EGF_1; 2.					
DR	PROSITE; PS01186; EGF_2; 1.					
DR	PROSITE; PS01187; EGF_Ca; 1.					
DR	PROSITE; PS01241; Link; 2.					
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.					
DR	PROSITE; PS00041; C_TYPE_LECTIN_2; 1.					
KW	Glycoprotein; Proteoglycan; Lectin; Extracellular matrix; Sushi;					
KW	Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;					
KW	Hyaluronic acid; Alternative splicing.					
FT	SIGNAL	1	20			POTENTIAL.
FT	CHAIN	21	3358			VERSICAN CORE PROTEIN.
FT	DOMAIN	37	137			IG-LIKE V-TYPE DOMAIN.
FT	DOMAIN	167	244			LINK 1.
FT	DOMAIN	265	346			LINK 2.
FT	DOMAIN	348	1308			GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
FT	DOMAIN	1309	3052			GAG-BETA.
FT	DOMAIN	3090	3126			EGF-LIKE 1.
FT	DOMAIN	3139	3253			EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	3258	3316			C-TYPE LECTIN.
FT	DISULFID	44	130			SUSHI.
FT	DISULFID	172	243			BY SIMILARITY.
FT	DISULFID	196	217			BY SIMILARITY.
FT	DISULFID	270	333			BY SIMILARITY.
FT	DISULFID	294	315			BY SIMILARITY.
FT	DISULFID	3056	3067			BY SIMILARITY.
FT	DISULFID	3061	3076			BY SIMILARITY.
FT	DISULFID	3078	3087			BY SIMILARITY.
FT	DISULFID	3094	3105			BY SIMILARITY.
FT	DISULFID	3099	3114			BY SIMILARITY.
FT	DISULFID	3116	3125			BY SIMILARITY.
FT	DISULFID	3132	3143			BY SIMILARITY.
FT	DISULFID	3160	3252			BY SIMILARITY.
FT	DISULFID	3228	3244			BY SIMILARITY.
FT	DISULFID	3259	3302			BY SIMILARITY.
FT	DISULFID	3288	3315			BY SIMILARITY.
FT	CARBOHYD	57	57			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	330	330			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	351	351			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	441	441			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	807	807			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	914	914			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	951	951			N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1305	1305			N-LINKED (GLCNAC. .) (POTENTIAL).

[4] SEQUENCE OF 2711-3396 FROM N.A.
 RC TISSUE="Lung fibroblast";
 RX MEDLINE-88007514; PubMed-2820964;
 RA Krusius T.; Gehlsen K.R.; Ruoslahti E.;
 RT "A fibroblast chondroitin sulfate proteoglycan core protein contains
 lectin-like and growth factor-like sequences.";
 RL J. Biol. Chem. 262:13120-13125(1987).
 [5] SEQUENCE OF 251-347 FROM N.A.
 RP MEDLINE-93122792; PubMed-1478664;
 RX Iozzo R.V.; Naso M.F.; Cannizzaro L.A.; Wasmuth J.J.,
 RA McPherson J.D.;
 RT "Mapping of the versican proteoglycan gene (CSPG2) to the long arm of
 human chromosome 5 (5q12-5q14).";
 RL Genomics 14:845-851(1992).
 [6] SEQUENCE FROM N.A. (ISOFORM V3).
 RC TISSUE=Brain;
 RX MEDLINE-95181355; PubMed-7876137;
 RA Zako M.; Shinomura T.; Ujita M.; Ito K.; Kimata K.;
 RT "Expression of PG-M(V3), an alternatively spliced form of PG-M
 without a chondroitin sulfate attachment in region in mouse and human
 tissues.";
 RT J. Biol. Chem. 270:3914-3918(1995).
 [7] SEQUENCE OF 3333-3396 FROM N.A. (ISOFORM VINT).
 RC TISSUE=Aortic smooth muscle;
 RX MEDLINE-99327053; PubMed-10397680;
 RA Lemire J.M.; Braun K.R.; Maurel P.; Kaplan E.D.; Schwartz S.M.,
 RA Wight T.N.;
 RT "Versican/PG-M isoforms in vascular smooth muscle cells.";
 RL Arterioscler. Thromb. Vasc. Biol. 19:1630-1639(1999).
 [8] PARTIAL SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE-89174663; PubMed-2466833;
 RA Perides G.; Lane W.S.; Andrews D.; Dahl D.; Bignami A.;
 RT "Solation and partial characterization of a glial
 hyaluronate-binding protein";
 RL J. Biol. Chem. 264:5981-5987(1989).
 [9] TISSUE SPECIFICITY OF ISOFORMS.
 RX MEDLINE-96213482; PubMed-8627343;
 RA Paulus W.; Baur J.; Dours-Zimmermann M.T.; Zimmermann D.R.;
 RT "Differential expression of versican isoforms in brain tumors.";
 RL J. Neuropathol. Exp. Neurol. 55:528-533(1996).
 CC -1- FUNCTION: May play a role in intercellular signaling and in
 connecting cells with the extracellular matrix. May take part in
 the regulation of cell motility, growth and differentiation. Binds
 hyaluronic acid.
 CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
 CC -1- ALTERNATIVE PRODUCTS: At least 5 isoforms; v0 (shown here), v1,
 v2, v3 and vint; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: cerebral white matter. v0 and v1 is expressed
 in normal brain, gliomas, medulloblastomas, schwannomas,
 neurofibromas, and meningiomas; v2 is restricted to normal brain
 and gliomas; v3 is found in all these tissues except
 medulloblastomas.
 CC -1- DEVELOPMENTAL STAGE: Disappears after the cartilage development.
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 LINK DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC CC DR EMBL; U16306; AAA65018.1;
 CC DR EMBL; X15998; CAA34128.1;
 CC DR EMBL; S52488; AAB24878.1;
 CC DR EMBL; U26555; AAA67565.1;
 CC DR EMBL; D32039; BAA06801.1;
 CC DR EMBL; J02814; AAA36437.1;
 CC DR EMBL; AF084545; AAD48545.1;
 CC DR PIR; S06014; S06014;
 CC DR PIR; A29348; A29348;
 CC DR PIR; A30358; A30358;
 CC DR HSSP; P01132; 1EGF.
 CC DR Genew; HGNC:2464; CSPG2.
 CC DR MIM; 118661;
 CC DR InterPro; IPR000152; Asx_hydroxyl.
 CC DR InterPro; IPR000561; EGF-like.
 CC DR InterPro; IPR000742; EGF_2.
 CC DR InterPro; IPR001881; EGF_Ca.
 CC DR InterPro; IPR001438; EGF_II.
 CC DR InterPro; IPR003599; Ig.
 CC DR InterPro; IPR003006; Ig_MHC.
 CC DR InterPro; IPR001304; Lectin_C.
 CC DR InterPro; IPR000538; Link.
 CC DR InterPro; IPR000436; Sushi_SCR_CCP.
 CC DR Pfam; PF00008; EGF_2.
 CC DR Pfam; PF00047; Ig_1.
 CC DR Pfam; PF00059; lectin_C; 1.
 CC DR Pfam; PF00084; sushi_1.
 CC DR Pfam; PF00193; Xlink_2.
 CC DR PRINTS; PRO0010; EGFBLOOD.
 CC DR ProDom; PD000918; Link_2.
 CC DR SMART; SM00032; CCP; 1.
 CC DR SMART; SM00034; CLECT; 1.
 CC DR SMART; SM000179; EGF_Ca; 1.
 CC DR SMART; SM00001; EGF-like; 1.
 CC DR SMART; SM00409; Ig_1.
 CC DR SMART; SM00445; Link; 2.
 CC DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 CC DR PROSITE; PS00022; EGF_1; 2.
 CC DR PROSITE; PS01186; EGF_2; 1.
 CC DR PROSITE; PS01187; EGF_Ca; 1.
 CC DR PROSITE; PS01241; Link; 2.
 CC DR PROSITE; PS00615; C_TYPELECTIN_1; 1.
 CC DR PROSITE; PS50041; C_TYPELECTIN_2; 1.
 CC KW Glycoprotein; Proteoglycan; Lectin_Extracellular matrix; Sushi;
 CC KW Signal; Repeat; EGF-like domain; Calcium; Immunoglobulin domain;
 CC KW Hyaluronic acid; Alternative splicing.
 CC FT SIGNAL 1 20 POTENTIAL.
 CC FT CHAIN 21 3396 VERSICAN CORE PROTEIN.
 CC FT DOMAIN 37 137 IG-LIKE V-TYPE DOMAIN.
 CC FT DOMAIN 167 244 LINK 1.
 CC FT DOMAIN 265 346 LINK 2.
 CC FT DOMAIN 348 1335 GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT DOMAIN).
 CC FT DOMAIN 1336 3089 GAG-BETA.
 CC FT DOMAIN 3089 3125 EGF-LIKE 1.
 CC FT DOMAIN 3127 3163 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 CC FT DOMAIN 3176 3290 C-TYPE LECTIN.
 CC FT DOMAIN 3295 3353 SUSHI.
 CC FT DISULFID 44 130 BY SIMILARITY.
 CC FT DISULFID 172 243 BY SIMILARITY.
 CC FT DISULFID 196 217 BY SIMILARITY.
 CC FT DISULFID 270 345 BY SIMILARITY.
 CC FT DISULFID 294 315 BY SIMILARITY.
 CC FT DISULFID 3093 3104 BY SIMILARITY.
 CC FT DISULFID 3098 3113 BY SIMILARITY.
 CC FT DISULFID 3115 3124 BY SIMILARITY.
 CC FT DISULFID 3131 3142 BY SIMILARITY.
 CC FT DISULFID 3136 3151 BY SIMILARITY.
 CC FT DISULFID 3153 3162 BY SIMILARITY.
 CC FT DISULFID 3169 3180 BY SIMILARITY.
 CC FT DISULFID 3197 3289 BY SIMILARITY.

FT	DISULFID	3265	3281	BY SIMILARITY.	
FT	DISULFID	3296	3339	BY SIMILARITY.	
FT	DISULFID	3325	3352	BY SIMILARITY.	
FT	CARBOHYD	57	57	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	330	330	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	615	615	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	782	782	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	809	809	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1332	1332	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1398	1398	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1442	1442	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1468	1468	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1663	1663	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1898	1898	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	2179	2179	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	2272	2272	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	2280	2280	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	2360	2360	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	2385	2385	N-LINKED (GLCNAC. . .) (POTENTIAL).	
Query	WIGL	76	WIGLRRREEKQSNSTACQDLYAWTDGSSISQFRNWYVDEP---	Score 174; DB 1; Length 3396;	8;
Best Local Similarity	28.5%	Pred. No.	3.9e-07;		
Matches	47;	Conservative	23;	Mismatches 55; Indels 40; Gaps	
Qy	QRPCYKVIFYFHDTSRRLNFEAKEACRRDGQQLVSIESEDEQKLIEKFIEENLLPSDGDF-	17	75		
Db	QGQCYK--YF--	3177	AHRRRTWDAEERECLQGAHLTSILSHEEQMFVNRV-----	GHDYQ 3224	
Qy	NDKMFEHDFRWTDGSTLQYENWRPNQPDSSFFSAGEDCVVIIWHENG--	3225	3274		
Qy	AGIGGPMFQWNDDRCNMKNNFICKYS---	132	DEKPAVPSREAEGE 172		
Db	QWNNDVPCNYHILTYTCKKGTVACGGBPVWENAKTECK	3275	3310		

RESULT 10

PGCV_CHICK

ID	PGCV	CHICK	STANDARD:	PRT:	3562 AA.
AC	Q90953;	Q90945;			
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Versican core protein precursor (Large fibroblast proteoglycan (Chondroitin sulfate proteoglycan core protein 2) (PG-M)).				
GN	CPSG2.				
OS	Gallus gallus (Chicken).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.				
OC	NCBI_TaxID=9031;				
RN	RP	SEQUENCE FROM N.A.			
RC	STRAIN-White Leghorn; TISSUE-Limb bud;				
RX	MEDLINE-93300846; PubMed-8314802;				
RA	Shinomura T., Nishida Y., Ito K., Kimata K.;				
RT	"cDNA cloning of PG-M, a large chondroitin sulfate proteoglycan expressed during chondrogenesis in chick limb buds. Alternative spliced multiforms of PG-M and their relationships to versican.";				
RL	J. Biol. Chem. 268:14461-14469 (1993).				
-i-	FUNCTION: May play a role in intercellular signaling and in connecting cells with the extracellular matrix. May take part in the regulation of cell motility, growth and differentiation. Binds hyaluronic acid.				
-i-	SUBCELLULAR LOCATION: Secreted; extracellular matrix.				
-i-	ALTERNATIVE PRODUCTS: At least 2 isoforms; V0 (shown here) and V1; are produced by alternative splicing.				
-i-	TISSUE SPECIFICITY: Prechondrogenic condensation area of developing limb buds.				
-i-	DEVELOPMENTAL STAGE: Disappears after the cartilage development (by similarity).				
-i-	SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.				

PT CARBOHYD 235 235 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 529 529 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 709 709 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 948 948 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1409 1409 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1479 1479 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1523 1523 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1530 1530 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1625 1625 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1751 1751 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 1988 1988 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2088 2088 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2089 2089 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2507 2507 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2642 2642 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2679 2679 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2748 2748 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 2762 2762 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 3069 3069 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 3194 3194 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 3232 3232 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT CARBOHYD 3545 3545 N-LINKED (GLCNAC. .) (POTENTIAL).
 PT VARSPLIC 485 1411 MISSING (IN ISOFORM V1).
 SQ SEQUENCE 3562 AA; 388078 MW; 9BC566E88C1602D2 CRC64;

Query Match 15.3%; Score 171; DB 1; Length 3562;
 Best Local Similarity 28.5%; Pred. No. 7.6e-07;
 Matches 47; Conservative 23; Mismatches 55; Indels 40; Gaps 8;

Qy 17 QRPCYKVIVFHDTSSRLNFEAKEACRRDGQLVSISSDEQKLIEKFIEENLLPSDGDF- 75
 Db 3342 QGQCYK--YF---AHRRTWDTAERECLQGAHLSIISHEEQFVNRI-----GHDYQ 3389
 Qy 76 WIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYVDEP----SCGSEVCVVMYHQPSAP 131
 Db 3390 WIGL-----NDKMFERDFRWTDGSPQYENWRPNQPDSSFSSAGEDCVVIIWHENG-- 3439
 Qy 132 AGIGGPPMFWQNDRCNMKNNFICKYS---DEKPAVPSREAGEE 172
 Db 3440 -----QWNDDVPCNYHLLTYtCKKGTVACGQPPVVENAKTFGK 3475

RESULT 11
 CDD93_RAT
 ID CDD93_RAT STANDARD; PRT; 643 AA.
 AC Q9ET61; Q9J1Z6;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Complement component C1q receptor precursor (Complement component 1, q
 DE subcomponent, receptor 1) (C1QR(p) (C1q/MBL/SPA receptor)
 DE (CDD93 antigen) (Cell surface antigen AA4).
 GN C1QR1 OR CD93 OR C1QRP.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID-10116;
 RN [1] SEQUENCE FROM N.A.
 RP STRAIN=PVG; TISSUE=Natural killer cells;
 RX MEDLINE=20545218; PubMed=11093152;
 RA Lovix G., Vagge J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
 RT "Characterization and molecular cloning of rat C1qRP, a receptor on NK
 cells.";
 RT Eur. J. Immunol. 30:3355-3362(2000).
 RN [2] SEQUENCE FROM N.A.
 RP STRAIN=Wistar; TISSUE=Lung;
 RX MEDLINE=20507883; PubMed=10934210;
 RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
 RT "Molecular and cellular properties of the rat AA4 antigen, a C-type
 lectin-like receptor with structural homology to thrombomodulin.";
 RT E->K (IN REF. 2).

RL J. Biol. Chem. 275:34382-34392(2000).
 CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for
 CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
 CC protein A (SPA). May mediate the enhancement of phagocytosis in
 CC monocytes and macrophages upon interaction with soluble defense
 CC collagens. May play a role in intercellular adhesion.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
 CC heart. Expressed at lower level in brain, thymus, liver, spleen,
 CC intestine, kidney, adrenal gland, muscle and testis. Expressed on
 CC endothelial cells, platelets, undifferentiated monocytes and
 CC circulating natural killer cells.
 CC -!- PTM: N- and O-glycosylated (By similarity).
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 DR EMBL; AF136537; AAG01572.1; -.
 DR EMBL; AF160978; AAF80402.1; -.
 DR HSSP; P35555; 1EMN.
 DR InterPro; IPR000152; Asx_hydroxy1.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001304; Lectin_C.
 PFam; PF00008; EGF; 5.
 PFam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 5.
 DR SMART; SM00179; EGF_Ca; 5.
 DR SMART; SM00001; EGF_1like; 2.
 DR PROSITE; PS00010; ASX_HYDROXYL; 3.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 DR PROSITE; PS01186; EGF_2; 3.
 DR PROSITE; PS01187; EGF_Ca; 3.
 KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;
 KW Repeat; Lectin; Lectin.
 KW SIGNAL 1 23
 FT CHAIN 24 643
 FT DOMAIN 24 571
 FT TRANSMEM 572 592
 FT DOMAIN 593 643
 FT DOMAIN 31 173
 FT DOMAIN 257 298
 FT DOMAIN 299 341
 FT DOMAIN 342 381
 FT DOMAIN 382 423
 FT DOMAIN 424 462
 FT DISULFID 261 272
 FT DISULFID 268 282
 FT DISULFID 284 297
 FT DISULFID 303 314
 FT DISULFID 308 325
 FT DISULFID 327 340
 FT DISULFID 346 355
 FT DISULFID 351 364
 FT DISULFID 366 380
 FT DISULFID 386 397
 FT DISULFID 393 406
 FT DISULFID 408 422
 FT DISULFID 428 437
 FT DISULFID 433 446
 FT DISULFID 448 461
 FT CARBOHYD 322 322
 FT CARBOHYD 498 498
 FT CONFLICT 417 417

SO	SEQUENCE	643 AA;	68781 MW;	9AE4C933AD943DB6 CRC64;	RL	FEBS Lett. 176:37-42(1984).
	Query Match	14.8%;	Score 165;	DB 1;	RP	PARTIAL SEQUENCE.
	Best Local Similarity	25.2%;	Pred. No. 3.5e-07;		RX	MEDLINE-87005253; PubMed-3530809;
	Matches 52;	Conservative 35;	Mismatches 73;	Indels 46;	RA	Perin J.-P.; Bonnet F.; Jolles P.;
				Gaps 10;	RT	"Structural relationship between link proteins and proteoglycan monomers.";
Oy	3	GRLLSGQP-----	YCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRRDGQLVSEED 56		RL	FEBS Lett. 206:73-77(1986).
Db	15	GQLWAGAAADSEAVVCEG--	TACYTAHW----GKLSSAAEAQHRCNENGNLATVKSEE 66	CC	MATRIX OF CARTILAGEOUS TISSUES. A MAJOR FUNCTION OF EXTRACELLULAR MATRIX IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.	
Oy	57	EQKLIEKFLENL----PSD----GDFWIGLRRREEKQSNSTACQDLYAWT-DGSISQFRN 108		CC	-1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).	
Db	67	EARHVQEALAAQLLKTAKPSETKIGKFWIGLQREKGKCTYHDLPMKGFSWVGGEDTTYSN 126		CC	-1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE C-TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2 AND G3.	
Oy	109	WY-VDEPSCGSECVVMMY-----HQPAPAGIGGPPMFQWNNDRC-----NMKNNF 153		CC	-1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO N-LINKED AND O-LINKED (ABOUT 40) OLIGOSACCHARIDES.	
Db	127	WYKASKSSCISKRCVSLIDLSLKPHPSHLP-----KWHESPCGTPDAPGNSIEGF 177		CC	-1- PTM: THE KERATAN SULFATE CONTENTS DIFFER CONSIDERABLY BETWEEN ADULT AND FETAL BOVINE PROTEOGLYCANS.	
Oy	154	ICKYSDEKPAVPSREAEGETEELTTP 179		CC	-1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.	
Db	178	LCKFNFKGMCSPPLALGGPGQLTYTTP 203		CC	-1- SIMILARITY: CONTAINS 4 LINK DOMAINS.	
RESULT 1.2						
PGCA_BOVIN						
ID	PGCA_BOVIN	STANDARD;	PRT;	2364 AA.	DR	EMBL; U76615; AAB38524.1; -
AC	P13608; Q28159; P79117;				DR	EMBL; L07053; -; NOT_ANNOTATED_CDS.
DT	01-JAN-1990 (Rel. 13, Created)				PIR	A27752; A27752.
DT	01-NOV-1997 (Rel. 35, Last sequence update)				DR	PIR; A29164; A29164.
DT	15-JUN-2002 (Rel. 41, Last annotation update)				DR	PIR; E29164; E29164.
DE	Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).				DR	PIR; G27751; G27751.
RN	SEQUENCE FROM N.A.				DR	HSSP; P08709; 1BF9.
RA	Hering T.M., Kollar J., Huynh T.D.;				DR	InterPro; IPR000152; Asx_hydroxyl.
RA	Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.				DR	InterPro; IPR000561; EGF-like.
RN	[2]	SEQUENCE OF 563-1056 FROM N.A.			DR	InterPro; IPR000742; EGF_2.
RP	SEQUENCE OF 563-1056 FROM N.A.				DR	InterPro; IPR001881; EGF_Ca.
RX	MEDLINE-89380219; PubMed-2528543;				DR	InterPro; IPR003599; Ig.
RA	Antonsson P., Heinegaard D., Oldberg A.;				DR	InterPro; IPR003006; Ig_MHC.
RT	"The keratan sulfate-enriched region of bovine cartilage proteoglycan consists of a consecutively repeated hexapeptide motif."				DR	InterPro; IPR001304; Lectin_C.
RT	deduced from a cDNA clone, contains numerous Ser-Gly sequences arranged in homologous repeats."				DR	InterPro; IPR000538; Link.
RL	J. Biol. Chem. 264:16170-16173(1989).				DR	InterPro; IPR003324; SGXXSG.
RN	[3]	SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.			DR	InterPro; IPR000436; Sushi_SCR_CCP.
RP	SEQUENCE OF 1609-2113 AND 2151-2364 FROM N.A.				DR	Pfam; PF00008; EGF; 1.
RX	MEDLINE-87270630; PubMed-311460;				DR	Pfam; PF00047; Ig; 1.
RA	Oldberg A., Antonsson P., Heinegaard D.;				DR	Pfam; PF00059; lectin_C; 1.
RT	"The partial amino acid sequence of bovine cartilage proteoglycan deduced from a cDNA clone, contains numerous Ser-Gly sequences arranged in homologous repeats."				DR	Pfam; PF00084; sushi; 1.
RT	arranged in homologous repeats."				DR	Pfam; PF0193; xLink; 4.
RL	Biochem. J. 243:255-259(1987).				DR	ProDom; PD000918; Link; 4.
RN	[4]	SEQUENCE OF 2114-2150 FROM N.A.			DR	SMART; SM00032; CCP; 1.
RC	TISSUE=Cartilage;				DR	SMART; SM00034; CLECT; 1.
RX	MEDLINE-93352525; PubMed-8349621;				DR	SMART; SM00179; EGF_Ca; 1.
RA	Fueloep C., Walcz E., Valyon M., Grant T.T.;				DR	SMART; SM00409; Ig; 1.
RT	"Expression of alternatively spliced epidermal growth factor-like domains in aggrecans of different species. Evidence for a novel module."				DR	SMART; SM00445; Link; 4.
RT	RL. J. Biol. Chem. 268:17377-17383(1993).				DR	PROSITE; PS00010; ASX_HYDROXYL; 1.
RN	[5]	PARTIAL SEQUENCE.			DR	PROSITE; PS00022; EGF_1; 1.
RP	PARTIAL SEQUENCE.				DR	
RX	MEDLINE-85027710; PubMed-6489519;				DR	
RA	Perin J.-P., Bonnet F., Jolles J., Jolles P.;				DR	
RT	"Sequence data concerning the protein core of the cartilage proteoglycan monomers. Characterization of a sequence allowing the synthesis of an oligonucleotide probe."				DR	

DR	GeneID: HGNC:3119; AGC1.	FT	CARBOHYD	657	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	MMID: 155760;	FT	CARBOHYD	737	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	InterPro; IPR000561; EGF-like.	FT	CARBOHYD	1898	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	InterPro; IPR003599; Ig.	FT	VARSPLIC	2163	MISSING (IN ISOFORM 2 AND ISOFORM 3).
DR	InterPro; IPR003006; Ig_MHC.	FT	VARSPLIC	2330	MISSING (IN ISOFORM 3).
DR	InterPro; IPR001304; Lectin_C.	FT	CONFLICT	766	E->A (IN REF. 4).
DR	InterPro; IPR000538; Link.	FT	CONFLICT	847	E->V (IN REF. 4).
DR	InterPro; IPR003324; SGXXSG.	FT	CONFLICT	1928	E->A (IN REF. 2).
DR	InterPro; IPR000436; Sushi_SCR_CCP.	FT	CONFLICT	1964	I->V (IN REF. 2 AND 3).
DR	Pfam; PF00008; EGF; 1.	FT	CONFLICT	2070	P->A (IN REF. 3).
DR	Pfam; PF00047; Ig; 1.	FT	CONFLICT	2391	A->P (IN REF. 2 AND 3).
DR	Pfam; PF00059; lectin_C; 1.	SQ	SEQUENCE	2415	AA: 250191 MW: 1288937E1B98C6B6 CRC64;
DR	Pfam; PF00084; sushi; 1.				
DR	Pfam; PF00193; Xlink; 4.				
DR	Pfam; PF02339; SGXXSG; 71.				
DR	ProDom; PD000918; Link; 4.				
DR	SMART; SM00032; CCP; 1.				
DR	SMART; SM00034; CLECT; 1.				
DR	SMART; SM00181; EGF; 1.				
DR	SMART; SM00409; Ig; 1.				
DR	SMART; SM00445; Link; 4.				
DR	PROSITE; PS00022; EGF; 1;				
DR	PROSITE; PS01186; EGF; 2;				
DR	PROSITE; PS00290; Ig_MHC; 1.				
DR	PROSITE; PS01241; Link; 3.				
DR	PROSITE; PS00615; C_TYPE_LECTIN_1; 1.				
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.				
KW	Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain; Alternative splicing; Repeat; Immunoglobulin domain.				
FT	SIGNAL	1	19	POTENTIAL.	
FT	CHAIN	20	2415	AGGRECAN CORE PROTEIN.	
FT	DOMAIN	44	140	IG-LIKE V-TYPE DOMAIN.	
FT	DOMAIN	170	247	LINK 1.	
FT	DOMAIN	268	349	LINK 2.	
FT	DOMAIN	495	572	LINK 3.	
FT	DOMAIN	593	673	LINK 4.	
FT	DOMAIN	2164	2199	EGF-LIKE.	
FT	DOMAIN	2201	2327	C-TYPE LECTIN.	
FT	DOMAIN	2331	2389	SUSHI.	
FT	DOMAIN	48	141	G1-A.	
FT	DOMAIN	152	247	G1-B.	
FT	DOMAIN	253	349	G1-B'.	
FT	DOMAIN	477	571	G2-B.	
FT	DOMAIN	578	672	G2-B'.	
FT	DOMAIN	676	848	KS.	
FT	DOMAIN	772	844	12 X APPROXIMATE TANDEM REPEATS.	
FT	DOMAIN	851	1497	CS-1.	
FT	DOMAIN	941	1497	29 X APPROXIMATE TANDEM REPEATS.	
FT	DOMAIN	1498	2162	CS-2.	
FT	DISULFID	2163	2415	G3.	
FT	DISULFID	51	133	BY SIMILARITY.	
FT	DISULFID	175	246	BY SIMILARITY.	
FT	DISULFID	199	220	BY SIMILARITY.	
FT	DISULFID	273	348	BY SIMILARITY.	
FT	DISULFID	297	318	BY SIMILARITY.	
FT	DISULFID	500	571	BY SIMILARITY.	
FT	DISULFID	524	545	BY SIMILARITY.	
FT	DISULFID	598	672	BY SIMILARITY.	
FT	DISULFID	621	642	BY SIMILARITY.	
FT	DISULFID	2168	2178	BY SIMILARITY.	
FT	DISULFID	2173	2187	BY SIMILARITY.	
FT	DISULFID	2189	2198	BY SIMILARITY.	
FT	DISULFID	2205	2216	BY SIMILARITY.	
FT	DISULFID	2233	2325	BY SIMILARITY.	
FT	DISULFID	2301	2317	BY SIMILARITY.	
FT	DISULFID	2332	2375	BY SIMILARITY.	
FT	DISULFID	2361	2388	BY SIMILARITY.	
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	239	239	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	387	387	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	434	434	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	602	602	N-LINKED (GLCNAC. . .) (POTENTIAL).	
Query Match	9	QPVCRGG---TQRPCYKVYFHDTSRRLNFEEAKEACRDRGGQLVSVIESEDEQKLIEKFI	65	Score 158.5; DB 1; Length 2415;	
Best Local Similarity	29.1%	Pred. No. 6.3e-06;			
Matches	52;	Conservative 22; Mismatches 62;	Indels 43;	Gaps 11;	
QY	66	ENLLPSDGFDF-WIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNYYVDEPS---CGSEVC	121	14.2%; DB 1; Length 2415;	
Db	2202	QEYCEEGLWNKYQGHCYR---HFPD---RETWVDAERRCREEQQSHLSSIVTPEQ	65	14.2%; DB 1; Length 2415;	
QY	2253	NN---NAQDQYWIGL---NDRTIEGDFRWSDGHPMQFENWRPNQDPDNFFAAGEDC	2301	14.2%; DB 1; Length 2415;	
Db	2302	VVM-YHOPSAPIAGIGGPMFQWNDDRCNMKNNFICKYS---DEKPAPVPSREAEGEETE	175	14.2%; DB 1; Length 2415;	
RESULT 14					
PGCA_CANFA					
ID	PGCA_CANFA	STANDARD;	PRT;	2333 AA.	
AC	Q28343; Q28310;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	16-OCT-2001 (Rel. 40, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Aggre can core protein precursor (Cartilage-specific proteoglycan core protein) (CSPCP).				
GN	AGC1.				
OS	Canis familiaris (Dog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
NCBI_TaxID	9615;				
[1]	SEQUENCE FROM N.A.				
RA	Giant T.T., Adams M.E., Kwok S.X.F., Huang D., Fullop C.;				
RA	"Complete coding sequence and deduced amino acid sequence of aggrecan of canine cartilage."				
RT	Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.				
RN	[2]				
RP	SEQUENCE OF 774-833 FROM N.A.				
RC	TISSUE-Cartilage;				
RX	MEDLINE-95128522; PubMed-7827755;				
RA	Barry F.P., Neame P.J., Sasse J., Pearson D.;				
RT	"Length variation in the keratan sulfate domain of mammalian aggrecan."				
RL	Matrix Biol. 14:323-328(1994).				
RN	[3]				
RP	SEQUENCE OF 1830-2333 FROM N.A.				
RA	Adams M.E., Kowk S.X.F., Huang D., Glant T.T.;				
RA	Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.				
RN	[4]				
RP	SEQUENCE OF 2082-2118 FROM N.A.				
RC	TISSUE-Cartilage;				
RX	MEDLINE-93352525; PubMed-8349621;				
RA	Fueloep C., Walcz E., Valyon M., Glant T.T.;				
RT	"Expression of alternatively spliced epidermal growth factor-like domains in aggrecans of different species. Evidence for a novel module."				
RL	J. Biol. Chem. 268:17377-17383(1993).				
CC	-1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR MATRIX OF CARTILAGENOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN				

CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.

CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix (By similarity).

CC -1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRIZE THE AMINO TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3, MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS) AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2 AND G3.

CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES (BY SIMILARITY).

CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.

CC -1- SIMILARITY: CONTAINS 4 LINK DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.

CC -1- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to licensee@isb-sib.ch).

CC DR EMBL; U65989; AAB06238.2; -.

CC DR EMBL; S74662; AAC60527.1; -.

CC DR EMBL; L07054; -; NOT_ANNOTATED_CDS.

CC DR HSSP; P08709; 1BF9.

CC DR InterPro; IPR000152; Asx_hydroxyl.

CC DR InterPro; IPR000561; EGF-like.

CC DR InterPro; IPR000742; EGF_2.

CC DR InterPro; IPR001881; EGF_Ca.

CC DR InterPro; IPR003599; Ig.

CC DR InterPro; IPR003006; Ig_MHC.

CC DR InterPro; IPR001304; Lectin_C.

CC DR InterPro; IPR000538; Link.

CC DR InterPro; IPR000436; Sushi_SCR_CCP.

CC DR Pfam; PF00008; EGF; 1.

CC DR Pfam; PF00047; 19; 1.

CC DR Pfam; PF00059; lectin_c; 1.

CC DR Pfam; PF00084; sushi; 1.

CC DR Pfam; PF00193; Xlink; 4.

CC DR Pfam; PF02339; SGXXSG; 70.

CC DR PRINTS; PRO1265; LINKMODULE.

CC DR SMART; SM00032; CCP; 1.

CC DR SMART; SM00034; CLECT; 1.

CC DR SMART; SM00179; EGF_Ca; 1.

CC DR SMART; SM00409; Ig; 1.

CC DR PROSITE; PS00010; ASX_HYDROXYL; 1.

CC DR PROSITE; PS00022; EGF_1; 1.

CC DR PROSITE; PS00290; Ig_MHC; 1.

CC DR PROSITE; PS01241; LINK; 4.

CC DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.

CC DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.

CC KW Glycoprotein; Cartilage; Proteoglycan; Lectin; Signal; Sushi; EGF-like domain; Repeat; Immunoglobulin domain.

CC FT SIGNAL 1 16 POTENTIAL.

CC FT CHAIN 17 2333 AGGREGAN CORE PROTEIN.

CC FT DOMAIN 44 140 IG-LIKE V-TYPE DOMAIN.

CC FT DOMAIN 170 247 LINK 1.

CC FT DOMAIN 349 816 LINK 2.

CC FT DOMAIN 513 590 LINK 3.

CC FT DOMAIN 611 692 LINK 4.

CC FT DOMAIN 2081 2117 EGF-LIKE, CALCIUM-BINDING (POTENTIAL).

CC C-TYPE LECTIN.

CC DOMAIN 2130 2249 2307 SUSHI.

CC DOMAIN 2130 2249 2307 G1-A.

CC DOMAIN 48 140 G2-B.

CC DOMAIN 152 247 G1-B.

CC DOMAIN 253 349 G1-B'.

CC DOMAIN 495 589 G2-B'.

CC DOMAIN 596 691 G2-B'.

CC DOMAIN 694 816 KS.

CC DOMAIN 819 1394 CS-1.

CC DOMAIN 1395 2079 CS-2.

CC DOMAIN 2080 2333 G3.

CC DISULFID 51 133 BY SIMILARITY.

CC DISULFID 175 246 BY SIMILARITY.

CC DISULFID 199 220 BY SIMILARITY.

CC DISULFID 273 348 BY SIMILARITY.

CC DISULFID 297 318 BY SIMILARITY.

CC DISULFID 518 589 BY SIMILARITY.

CC DISULFID 542 563 BY SIMILARITY.

CC DISULFID 616 691 BY SIMILARITY.

CC DISULFID 640 661 BY SIMILARITY.

CC DISULFID 2085 2096 BY SIMILARITY.

CC DISULFID 2090 2105 BY SIMILARITY.

CC DISULFID 2107 2116 BY SIMILARITY.

CC DISULFID 2123 2134 BY SIMILARITY.

CC DISULFID 2151 2243 BY SIMILARITY.

CC DISULFID 2219 2235 BY SIMILARITY.

CC DISULFID 2250 2293 BY SIMILARITY.

CC DISULFID 2279 2306 BY SIMILARITY.

CC CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 676 676 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 747 747 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SQ SEQUENCE 2333 AA; 240573 MW; 8B9ED78F3508B596 CRC64;

CC Query Match 13.9%.

CC Best Local Similarity 28.5%.

CC Matches 51; Conservative 24; Mismatches 61; Indels 43; Gaps 11;

Qy 9 QPVCRGG---TQRPCYKVIVFHDTSRRLNFEEAKEACRRDGGQLVSISEDEQKLIEKF1 65

Db 2120 QELCEEGWTKFQGHCYR--YFPD---RESWDAESRCRAQQSHLSSIVTPEEQ---EFV 2170

Qy 66 ENLLPSDGDF-WIGLRRREEKQNSNSTACQDLYAWTDGSIQFRNWYVDEPS---CGSEVC 121

Db 2171 NN---NAQDYQWIGL---NDRTIEGDFRWSDGHSIQLQENWRPNQPDNFFVSGEDC 2219

Qy 122 VVM-YHQPSAPAGIGGPMFQWNDDRCNMKNNFICKYS---DEKPAVPSREAEGEETE 175

Db 2220 VVMIWHEKG-----EWNDVPCNYYLPTCKKGTVACGDPPVVEHARTFGQKKD 2267

RESULT 15

LEM2_MOUSE STANDARD; PRT; 612 AA.

ID LEM2_MOUSE ID LEM2_MOUSE STANDARD; PRT; 612 AA.

AC Q00690; AC Q00690;

DT 01-APR-1993 (Rel. 25, Created)

DT 01-APR-1993 (Rel. 25, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2) (CD62E). SELE OR ELAM-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. [1]

OX NCBI_TAXID=10090;

RN SEQUENCE FROM N.A.

RX MEDLINE=92283265; PubMed=1375914;

RA Becker-Andre M.; van Huijsduijnen R.H.; Losberger C.; Whelan J.;
 RA Delamarre J.F.;
 RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
 RT structural and functional homologue of the human protein.";
 RT Eur. J. Biochem. 206:401-411(1992).
 [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE-92340571; PubMed-1378846;
 RX Weller A.; Isenmann S.; Vestweber D.;
 RA "Cloning of the mouse endothelial selectins. Expression of both
 RT and P-selectin is inducible by tumor necrosis factor alpha.";
 RL J. Biol. Chem. 267:15176-15183(1992).
 CC -i- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES
 CC POLYGLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
 CC GLYCOLIPIDS).
 CC -i- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -i- SIMILARITY: BELONGS TO THE SELECTIN/LECAM FAMILY.
 CC -i- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -i- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 CC -i- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a col-
 CC laboration between the Swiss Institute of Bioinformatics and the EMBL ou-
 CC the European Bioinformatics Institute. There are no restrictions
 CC on use by non-profit institutions as long as its content is not
 CC modified and this statement is not removed. Usage by and for
 CC entities requires a license agreement (see <http://www.isb-sib.ch>)
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; M80778; AAA37547.1; -.
 DR EMBL; M87862; AAA37577.1; ALT_INIT.
 DR HSSP; P16581; 1KJA.
 DR MGD; MGI:98278; Sele.
 DR InterPro; IPR000561; EGF-like.
 DR InterPro; IPR001304; Lectin_C.
 DR InterPro; IPR002396; Selectin.
 DR InterPro; IPR000436; Sushi_SCR_CCP.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00059; lectin_C; 1.
 DR Pfam; PF00084; sushi; 6.
 DR PRINTS; PR00343; SELECTIN.
 DR SMART; SM00032; CCP; 6.
 DR SMART; SM00034; CLECT; 1.
 DR SMART; SM00181; EGF; 1.
 DR PROSITE; PS00022; EGF_1; 1.
 DR PROSITE; PS01186; EGF_2; 1.
 DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
 DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
 DR Cell adhesion; Transmembrane; Glycoprotein; Lectin.
 KW Selectin; Signal; Sushi; Repeat.
 FT SIGNAL 1 21 BY SIMILARITY.
 FT CHAIN 22 612 E-SELECTIN.
 FT DOMAIN 22 557 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 558 579 POTENTIAL.
 FT DOMAIN 580 612 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 38 138 C-TYPE LECTIN (SHORT FORM).
 FT DOMAIN 139 175 EGF-LIKE.
 FT DOMAIN 179 239 SUSHI 1.
 FT DOMAIN 242 301 SUSHI 2.
 FT DOMAIN 304 364 SUSHI 3.
 FT DOMAIN 367 427 SUSHI 4.
 FT DOMAIN 430 490 SUSHI 5.
 FT DOMAIN 493 549 SUSHI 6.
 FT DISULFID 40 138 BY SIMILARITY.
 FT DISULFID 111 130 BY SIMILARITY.
 FT DISULFID 143 154 BY SIMILARITY.
 FT DISULFID 148 163 BY SIMILARITY.
 FT DISULFID 165 174 BY SIMILARITY.
 FT DISULFID 180 225 BY SIMILARITY.
 FT DISULFID 210 238 BY SIMILARITY.
 FT DISULFID 243 287 BY SIMILARITY.

GenCore version 5.1.4_p5_4578
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2003, 17:02:46 ; Search time 31 Seconds

(without alignments)
 1369.216 Million cell updates/sec

Title: US-09-887-855-5

Perfect score: 1115

Sequence: 1 ATGRLLSGQPVCRGGTQRPC.....EEDAKTFKESREAALNLY 206

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched:

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

1: sp_archea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rabbit:*

12: sp_rvirus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriap:*

17: sp_archeap:*

ALIGNMENTS

17 158 14.2 295 11 Q912W4 mus musculu

18 158 14.2 311 11 Q9D8V4 mus musculu

19 158 14.2 325 11 Q912X0 mus musculu

20 157 14.1 158 13 Q90W17 bungarus fa

21 156 14.0 339 6 Q95244 sus scrofa

22 155 13.9 1152 13 Q90WM2 xenopus lae

23 152.5 13.7 485 6 Q95LG3 odocoileus

24 152 13.6 158 13 Q90WI6 bungarus mu

25 150 13.5 197 6 Q28008 bos taurus

26 149.5 13.4 459 5 Q22136 caenorhabdi

27 149 13.4 158 13 Q90WI8 bungarus fa

28 148.5 13.3 195 5 Q27340 megabalanus

29 148.5 13.3 330 4 Q9NT67 homo sapien

30 148.5 13.3 911 4 Q9HBK4 homo sapien

31 148.5 13.3 911 4 Q96gW7 homo sapien

32 148.5 13.3 911 4 Q96fp7 homo sapien

33 148.5 13.3 911 4 Q8tbb9 homo sapien

34 148 13.3 253 5 Q20665 caenorhabdi

35 148 13.3 381 6 Q8SQB2 macaca mula

36 147 13.2 196 11 Q9EPW4 mus musculu

37 147 13.2 381 6 Q95LC6 macaca neme

38 147 13.2 404 6 Q95J96 macaca mula

39 146.5 13.1 152 13 Q9DG39 agkistrodon

40 146.5 13.1 162 5 Q25459 megabalanus

41 146.5 13.1 1031 5 Q8WSX2 dugesia tig

42 146 13.1 2109 13 P79787 gallus gall

43 145.5 13.0 134 5 Q9UB05 hydra atten

44 145.5 13.0 381 6 Q95LA8 macaca mula

45 145.5 13.0 404 6 Q95L98 pan troglod

RESULT 1

Q96NF3 PRELIMINARY;

AC Q96NF3 PRELIMINARY;

DT 01-DEC-2001 (TREMBLel. 19, Created)

DT 01-DEC-2001 (TREMBLel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLel. 21, Last annotation update)

DE CDNA FLJ30977 fis, Clone HHDPC2000095, highly similar to Cricetulus

DE griseus layillin mRNA.

DE Homo sapiens (Human).

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TAXID=9606;

[1] RNP

SEQUENCE FROM N.A.

RA Ishibashi T., Kanehori K., Yosida M., Watanabe S., Ishida S., Ono Y.,

RA Hotuta T., Hiraoaka S., Murakawa K., Takiguchi S., Kusano J.,

RA Watanabe M., Fujimori K., Tanai H., Ishida M., Yamashita H., Chiba Y.,

RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,

RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,

RA Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,

RA Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,

RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.,

RA "NEDO human DNA sequencing project".

RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AK055539; BAB70946.1;

DR InterPro; IPR001304; Lectin_C.

DR Pfam; PF00059; Lectin_c;

DR PROSITE; PS50041; C_TYPE_LECTIN_2;

1. SQ SEQUENCE 374 AA; 42280 MW; 8AE64E6BC9E56DCD CRC64;

2. Query %

Score 100.0%; Score 1115; DB 4;

Best Local Similarity 100.0%; Pred. No. 3.8e-98;

Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3. Description

4. %

5. Result No. Score Match Length DB ID Description

1 1115 100.0 374 4 Q96NF3 homologous protein

2 1115 100.0 374 4 Q8TAY8 homologous protein

3 111 99.6 374 4 Q96NC5 homologous protein

4 945.5 84.8 374 1 Q9Z209 cricetulus

5 574 51.5 273 1 Q8V131 mus musculus

6 185 16.6 1290 13 Q9W6E1 gallus gallus

7 178.5 16.0 1456 11 Q61830 mus musculus

8 177.5 15.9 1348 5 Q25199 hydra attenuata

9 176.5 15.8 1479 4 Q9Y5P9 homologous protein

10 176.5 15.8 1479 4 Q9UBG0 homologous protein

11 176 15.8 315 4 Q9UPK6 homologous protein

12 176 15.8 1321 4 Q14594 homologous protein

13 170.5 15.3 1479 11 Q64449 homologous protein

14 161.5 14.5 134 5 Q9XYX3 hydra magnipunctata

15 159.5 14.3 742 11 Q8VIF6 mus musculus

16 158.5 14.2 719 6 062623 bos taurus

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

1. Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

2. Sequence from N.A.

3. Best Local Similarity 100.0%; Pred. No. 3.8e-98; Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4. Sequence 374 AA; 42280 MW; 8AE64E6BC9E56DCD CRC64;

5. Result No. Score Match Length DB ID Description

1 1115 100.0 374 4 Q96NF3 homologous protein

2 1115 100.0 374 4 Q8TAY8 homologous protein

3 111 99.6 374 4 Q96NC5 homologous protein

4 945.5 84.8 374 1 Q9Z209 cricetulus

5 574 51.5 273 1 Q8V131 mus musculus

6 185 16.6 1290 13 Q9W6E1 gallus gallus

7 178.5 16.0 1456 11 Q61830 mus musculus

8 177.5 15.9 1348 5 Q25199 hydra attenuata

9 176.5 15.8 1479 4 Q9Y5P9 homologous protein

10 176.5 15.8 1479 4 Q9UBG0 homologous protein

11 176 15.8 315 4 Q9UPK6 homologous protein

12 176 15.8 1321 4 Q14594 homologous protein

13 170.5 15.3 1479 11 Q64449 homologous protein

14 161.5 14.5 134 5 Q9XYX3 hydra magnipunctata

15 159.5 14.3 742 11 Q8VIF6 mus musculus

16 158.5 14.2 719 6 062623 bos taurus

Qy	61 IEKFIENLLPSDGGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 120
Db	82 IEKFIENLLPSDGGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 141
Qy	121 CVVMYHQPSAPAGIGGPFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180
Db	142 CVVMYHQPSAPAGIGGPFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201
Qy	181 LPEETQEEDAKKTFKESREALNLAY 206
Db	202 LPEETQEEDAKKTFKESREALNLAY 227
RESULT 2	
Q8TAY8	PRELIMINARY; PRT; 374 AA.
ID Q8TAY8; AC Q8TAY8;	
DT 01-JUN-2002 (TREMBLrel. 21, Created)	
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE Similar to unnamed protein product.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	SEQUENCE FROM N.A.
RP	RC TISSUE=BRAIN;
RA Strausberg R.;	
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.	
DR EMBL; BC025407; AAH25407.1; -.	
SQ SEQUENCE 374 AA; 42312 MW; FC214E6BC9E578D9 CRC64;	
Query Match 100.0%; Score 1115; DB 4; Length 374;	
Best Local Similarity 100.0%; Pred. No. 3.8e-98;	
Matches 206; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Qy 1 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRQDGGQLVSIESEDEQKL 60	
Db 22 ATGRLLSGQPVCRGGTQRPCYKVIFYFHDTSRRLNFEEAKEACRQDGGQLVSIESEDEQKL 81	
Qy 61 IEKFIENLLPSDGGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 120	
Db 82 IEKFIENLLPSDGGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 141	
Qy 121 CVVMYHQPSAPAGIGGPFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180	
Db 142 CVVMYHQPSAPAGIGGPFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201	
Qy 181 LPEETQEEDAKKTFKESREALNLAY 206	
Db 202 LPEETQEEDAKKTFKESREALNLAY 227	
RESULT 3	
Q96NC5	PRELIMINARY; PRT; 374 AA.
ID Q96NC5; AC Q96NC5;	
DT 01-DEC-2001 (TREMBLrel. 19, Created)	
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)	
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	
DE CDNA FLJ31092 fis, clone IMR321000158, highly similar to Cricetulus griseus layillin.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	SEQUENCE FROM N.A.
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,	
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,	
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,	
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,	

RA	Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA	Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
RA	Kawakami B., Suzuki Y., Sugano S., Nagahashi K., Masuho Y., Nagai K.,
RA	Isogai T.;
RT	"NEDO human cDNA sequencing project.";
RL	Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AK055654; BAB70978.1; -.
DR	InterPro; IPR001304; Lectin_C.
DR	Pfam; PF00059; lectin_c; 1.
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR	SEQUENCE 374 AA; 42310 MW; CBF74E676E23BA16 CRC64;
SQ	
Query Match	99.6%; Score 1111; DB 4; Length 374;
Best Local Similarity	99.5%; Pred. No. 9.2e-98;
Matches	205; Conservative 0; Mismatches 1; Indels 0; Gaps
Qy	1 ATGRLLSQQPVCRGGTQRPCYKVIFYFHDTSRRLNFEAKEACRRDGQQLVSIEDEQKL 60
Db	22 ATGRLLSQQPVCRGGTQRPCYKVIFYFHDTSRRLNFEAKEACRRDGQQLVSIEDEQKL 81
Qy	61 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 120
Db	82 IEKFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEV 141
Qy	121 CVVMYHQPSAPAGIGGPFYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 180
Db	142 CVVMYHQPSAPAGIGGPFYMFQWNDDRCNMKNNFICKYSDEKPAVPSREAEGEETELTPV 201
Qy	181 LPEETQEEDAKKTFKESREAAALNAY 206
Db	202 LPEETQEEDTKKTFKESREAAALNAY 227
RESULT 4	
Q9Z209	PRELIMINARY; PRT; 374 AA.
ID	Q9Z209;
AC	Q9Z209;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE	Layillin.
OS	cricetulus griseus (Chinese hamster).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC	Cricetulus.
OX	NCBI_TaxID=10029;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Borowsky M.L., Hynes R.O.;
RT	"Layillin, a novel talin-binding transmembrane protein homologous with C-type lectins, is localized in membrane ruffles.";
RT	J. Cell Biol. 143:0-0(1998).
RL	EMBL; AF093673; AAC68695.1; -.
DR	HSSP; P06734; 1HLI.
DR	InterPro; IPR001304; Lectin_C.
DR	Pfam; PF00059; lectin_c; 1.
DR	SMART; SM00034; CLECT; 1.
DR	PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR	SEQUENCE 374 AA; 42435 MW; 298A8BA24FB04E1C CRC64;
SQ	
Query Match	84.8%; Score 945.5; DB 11; Length 374;
Best Local Similarity	85.9%; Pred. No. 5.5e-82;
Matches	176; Conservative 9; Mismatches 19; Indels 1; Gaps
Qy	3 GRLLSQQPVCRGGTQRPCYKVIFYFHDTSRRLNFEAKEACRRDGQQLVSIEDEQKLIE 62
Db	24 GRLLSQQQLVCRGGTRRPCYKVIFYFHDAFQRLNFEAKEACRRDGQQLVSIETEDEQRLIE 83
Qy	63 KFIENLLPSDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPSCGSEVCV 122
Db	84 KFIENLLASDGDFWIGLRRLEVQVNNTAACQDLYAWTDGSTSQFRNWWVDEPSCGSEVCV 143
Qy	123 VMYHQPSAPAGIGGPFYMFQWNDDRCNMKNNFICKYSDEKPA-VPSREAEGEETELTPV 181

RESULT 5

Q8VI31 PRELIMINARY; PRT; 273 AA.

ID Q8VI31; AC; DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE C-type lectin protein MT75.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TAXID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Weng L., Smits P., Hubner R., Wouteirs J., Merregaert J.;
RT "Mt75, a low expressed c-type lectin gene involving in
chondrogenesis."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AF311699; AAL50354.1; -.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR FROSTIE; PS50041; C_TYPE_LECTIN_2; 1.
KW Lectin.
SQ SEQUENCE 273 AA; 30303 MW; E052D933F244F4C7 CRC64;

Query Match 51.5%; Score 574; DB 11; Length 273;
Best Local Similarity 57.4%; Pred. No. 9.9e-47;
Matches 109; Conservative 26; Mismatches 41; Indels 14; Gaps 5;

QY 4 RLLSGQPVCRGGTQRPCYKVIYFHDTSRRLNFEEAKEACRRDGQLVSI
SIESEDEQKLIEK 63
DB 23 RVVSGQKVCFADVKHPCKMAYFHELSSRVSFQE^{ARLACESEGGVLLSLENEAEQK}LIES 82

QY 64 FIENLLP----SDGDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYD
EPSCGS 118
DB 83 MLQNLTKPGTG1SDGDFWIGLRLSGDGT-SGACPDLYQWSDGSSSQFRNWYD
EPSCGS 141

QY 119 EVCVVMYHOPSA^{PAGIGGPMFQWNDDRCNMKNNFICKYSDE}-KPAVPSRAE
GEETELT 177
DB 142 EKCVVVMYHOPTANPGLGGPYLYQWNDDRCNMKHYICTYEP^{EIHPTEPA}-
EKPILT 196

QY 178 TPVLPEETQE 187
DB 197 NQ--PEETHE 204

RESULT 6

Q9W6E1 PRELIMINARY; PRT; 1290 AA.

ID Q9W6E1; AC; DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Neurocan core protein.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TAXID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309833; PubMed=10851024;
RA Li H., Leung T.C., Hoffman S., Balsamo J., Lillian J., Harris N., Rits M., Chang G., Ezkowitz R.B.;

RT Coordinate Regulation of Cadherin and Integrin Function by the Chondroitin Sulfate Proteoglycan Neurocan.; J. Cell Biol. 149:1275-1288 (2000).
RL EMBL; AF116856; AAD24546.2; -.
DR HSSP; P08709; 1BF9.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000538; Link.
DR InterPro; IPR000436; Sushi_SCR_CCP.
PFam; PF00008; EGF; 2.
PFam; PF00047; Ig; 1.
PFam; PF00059; lectin_c; 1.
PFam; PF00193; Xlink; 2.
PRINTS; PR00010; EGF_BLOOD.
PRODOM; PD000918; Link; 2.
SMART; SM00032; CCP; 1.
SMART; SM00034; CLECT; 1.
SMART; SM00179; EGF_Ca; 1.
SMART; SM00001; EGF_like; 1.
SMART; SM00409; Ig; 1.
SMART; SM0045; Link; 2.
PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
PROSITE; PS00022; EGF_1; UNKNOWN_2.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS01187; EGF_Ca; 1.
PROSITE; PS01241; Link; 2.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat; SEQUENCE 1290 AA; 138877 MW; 182BD86D0E40BE78 CRC64;
SQ Score 16.6%; Length 1290;
Best Local Similarity 34.7%; Pred. No. 7.7e-09;
Matches 50; Conservative 18; Mismatches 42; Indels 34; Gaps 8;

Query Match 17 QRPCYKV1YFHDTSRRLNFEEAKEACRRDGQLVSI
SIESEDEQKLIEKFIENLLPSDGFW 76
DB 1064 QGHCYR--YF---SRRSWEDAERDCRRAGHLTSIHSQEEHGFINSF----GHENTW 1112

QY 77 IGLRRREEKQSNSTACQDLYAWTDGSISQFRNWYD
EPSCGS 118
DB 1113 IGLNDRIVEQD----FQWTDTNTGLOQYENWRENQPDNFFAGGEDC
VVLVSYE---- 1159

QY 133 GIGGPMFQWNDDRCNMKNNFICK 156
DB 1160 -IG----KWNNDVPCNYNLPYICK 1177

RESULT 7

Q61830 PRELIMINARY; PRT; 1456 AA.

ID Q61830; AC; DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Macrophage mannose receptor precursor.
GN MRC1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TAXID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=93043353; PubMed=1421407;
RA Harris N., Rits M., Chang G., Ezkowitz R.B.;

"Characterization of the murine macrophage mannose receptor.";
Blood 80:2363-2373 (1992).

[2]
RN SEQUENCE FROM N.A.

RC STRAIN-C57BL/6;

RA SUPER M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBBJ databases.

DR EMBL; Z11974; CAA78028.1; -.

DR HSSP; P22897; 1EGG.
MGD; MGI:97142; Mrc1.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR Pfam; PF00059; lectin_C; 8.

DR Pfam; PF00062; Ricin_B_lectin; 2.

DR PRINTS; PRO0013; FNTypeII.

DR PRODOM; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR PROSITE; PS000615; C_TYPELECTIN_1; 6.

DR PROSITE; PS500041; C_TYPELECTIN_2; 8.

DR PROSITE; PS500023; FIBRONECTIN_2; 1.

DR PROSITE; PS50231; Ricin_B_lectin; 1.

KW Receptor; Signal.

FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 1456 MACROPHAGE MANNOSE RECEPTOR.

SO SEQUENCE 1456 AA; 165065 MW; 4EBD3F1B8619A594 CRC64;

Query Match 16.0% Score 178.5; DB 11; Length 1456;

Best Local Similarity 25.9%; Pred. No. 3.7e-08;

Matches 53; Conservative 33; Mismatches 66; Indels 53; Gaps 8;

RESULT 9 Q9Y5P9

ID Q9Y5P9 PRELIMINARY; PRT; 1479 AA.

AC Q9Y5P9; PROSITE; PS000615; FN_Type_II; 12; Created

DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)

DT 01-NOV-1999 (TREMBLrel. 12; Last annotation update)

DT 01-JUN-2002 (TREMBLrel. 21; Last annotation update)

DE Endocytic receptor Endo180.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606; [1]

RN SEQUENCE FROM N.A. MEDLINE=20148849; PubMed=10683150;

RX RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;

RT "Endo180, an endocytic recycling glycoprotein related to the macrophage mannose receptor is expressed on fibroblasts, endothelial cells and macrophages and functions as a lectin receptor."

RL J. Cell Sci. 113:1021-1032(2000).

DR EMBL; AF134838; AAD30280.1; -.

DR HSSP; P02751; 2FN2.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR PRINTS; PR00013; FNTYPEII.

DR PRODom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR PROSITE; PS000615; C_TYPELECTIN_1; 3.

DR PROSITE; PS500041; C_TYPELECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

KW Receptor.

"Characterization of the murine macrophage mannose receptor.";
Blood 80:2363-2373 (1992).

[2]
RN SEQUENCE FROM N.A.

RC STRAIN-C57BL/6;

RA SUPER M.;
RL Submitted (APR-1992) to the EMBL/GenBank/DBBJ databases.

DR EMBL; Z11974; CAA78028.1; -.

DR HSSP; P22897; 1EGG.
MGD; MGI:97142; Mrc1.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR PRINTS; PR00013; FNTYPEII.

DR PRODom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; FN2; 1.

DR PROSITE; PS000615; C_TYPELECTIN_1; 3.

DR PROSITE; PS500041; C_TYPELECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

KW Receptor.

DR InterPro; IPR000719; Euk_Pkinase.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR001245; Tyr_Pkinase.

DR Pfam; PF00059; lectin_C; 4.

DR Pfam; PF00069; kinase; 1.

DR PRODom; PD000001; Euk_Pkinase; 2.

DR SMART; SM00219; TyrK_C; 1.

DR PROSITE; PS000615; C_TYPELECTIN_1; UNKNOWN_2.

DR PROSITE; PS500041; C_TYPELECTIN_2; 4.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.

KW ATP-binding; Kinase; Receptor; Transferase; Tyrosine-protein kinase.

SO SEQUENCE 1348 AA; 156916 MW; 77D2122093227FEF CRC64;

Query Match 15.9% Score 177.5; DB 5; Length 1348;

Best Local Similarity 28.0%; Pred. No. 4.2e-08;

Matches 47; Conservative 32; Mismatches 46; Indels 43; Gaps 8;

QY 6 LSGQPCV-RGGTQRPC----YKVYFHDTSRRLNFEAKEACRRDGGQLVSIESEDE 57

Db 412 LSHRFICKVKRATNEYCAEGWTSYRIVCYFISIESFDWFKSFSSCQNIGGNLNSIENQEE 471

QY 58 QKLIKEFIFIENLLPSGDD-FWIGLRR-----REEKQSSNSTACQDLYAWTDGSISQFRNWTY 110

Db 472 ---NRFIELDLIKNDKDYWIGLNIKTIWNDYKKNNKR-----FEWSDNTYTOFFNW1 518

QY 111 VDEP--SCGSEVCVMMYHQPSAPAGIGGPPMFQWNDRCNMKNNFICK 156

Db 519 TNQPDNNNGIESCVEMNN-----GWSDKECKVNLNGFICK 553

RESULT 9 Q9Y5P9

ID Q9Y5P9 PRELIMINARY; PRT; 1479 AA.

AC Q9Y5P9; PROSITE; PS000615; FN_Type_II; 12; Created

DT 01-NOV-1999 (TREMBLrel. 12; Last sequence update)

DT 01-NOV-1999 (TREMBLrel. 12; Last annotation update)

DT 01-JUN-2002 (TREMBLrel. 21; Last annotation update)

DE Endo180.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

NCBI_TaxID=9606; [1]

RN SEQUENCE FROM N.A. MEDLINE=20148849; PubMed=10683150;

RX RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;

RT "Endo180, an endocytic recycling glycoprotein related to the macrophage mannose receptor is expressed on fibroblasts, endothelial cells and macrophages and functions as a lectin receptor."

RL J. Cell Sci. 113:1021-1032(2000).

DR EMBL; AF134838; AAD30280.1; -.

DR HSSP; P02751; 2FN2.

DR InterPro; IPR000562; FN_Type_II.

DR InterPro; IPR001304; Lectin_C.

DR InterPro; IPR000566; Lipocln_cytFABP.

DR InterPro; IPR000772; Ricin_B_lectin.

DR Pfam; PF00040; fn2; 1.

DR PRINTS; PR00013; FNTYPEII.

DR PRODom; PD000995; FN_Type_II; 1.

DR SMART; SM00034; CLECT; 8.

DR SMART; SM00059; lectin_C; 8.

DR PROSITE; PS000615; C_TYPELECTIN_1; 3.

DR PROSITE; PS500041; C_TYPELECTIN_2; 8.

DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.

DR PROSITE; PS50231; RICIN_B_LECTIN; 1.

KW Receptor.

QY	7	SGOPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIESEDEQKLIKEKIE	66	7	SGOPVCRGGTQRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIESEDEQKLIKEKIE	66	
Db	385	SWQPF----QGHCYRL----QAEKRSWQESKKACLRRGGDLVSIHSMAALEFITKQIK	434	Db	385	SWQPF----QGHCYRL----QAEKRSWQESKKACLRRGGDLVSIHSMAALEFITKQIK	434
QY	67	NLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPS---CGSEVCVV	123	QY	67	NLLPSDGFDFWIGLRRREEKQSNSTACQDLYAWTDGSISQFRNWWVDEPS---CGSEVCVV	123
Db	435	QEVE---ELWIGL----NDLKLQMNEFSDGSSLVSFTHWHPFEPNFRDSLEDCVT	483	Db	435	QEVE---ELWIGL----NDLKLQMNEFSDGSSLVSFTHWHPFEPNFRDSLEDCVT	483
QY	124	MYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAE	171	QY	124	MYHQPSAPAGIGGPPMFQWNDDRCNMKNNFICKYSDEKPAVPSREAE	171
Db	484	IW----GPEG----RWNDSPCNQSLPSICKKAGOLSQGAAEEDHG	520	Db	484	IW----GPEG----RWNDSPCNQSLPSICKKAGOLSQGAAEEDHG	520
RESULT 11							
	Q9UPK6	PRELIMINARY;	PRT;	315	AA.		
	ID	Q9UPK6					
	AC	Q9UPK6;					
	DT	01-MAY-2000 (TREMBLrel.	13, Created)				
	DT	01-MAY-2000 (TREMBLrel.	13, Last sequence update)				
	DT	01-DEC-2001 (TREMBLrel.	19, Last annotation update)				
	DE	PGCN_HUMAN,	partial CDS (Fragment).				
	GN	NEUR.					
	OS	Homo sapiens (Human).					
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	OC	NCBI_TAXID=9606;					
	DT	01-MAY-2000 (TREMBLrel.	13, Last sequence update)				
	DT	01-MAY-2002 (TREMBLrel.	21, Last annotation update)				
	DE	Urokinase receptor-associated protein UPARAP.					
	GN	KIAA0709.					
	OS	Homo sapiens (Human).					
	OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
	OC	NCBI_TAXID=9606;					
	RN	[1]					
	RP	SEQUENCE FROM N.A.					
	RA	Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W., Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J., Danganan L., Poundstone P., Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T., Frankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S., Thomas P., Quan G., Kronmiller B., Arellano A., Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;					
	RA	"Sequence analysis of an -1 Mb region containing the MEF2B gene in [2]					
	RT	"Sequence analysis of an -1 Mb region containing the MEF2B gene in [2]"					
	RT	Submitted (JUL-1998) to the EMBL/GenBank/DBDJ databases.					
	RL	EMBL; AC005254; AAC25581.1; -.					
	DR	HSSP; P00740; 1EDM.					
	DR	InterPro; IPR000152; Asx_hydroxy1.					
	DR	InterPro; IPR000561; EGF-like.					
	DR	InterPro; IPR000742; EGF_2.					
	DR	InterPro; IPR001881; EGF_Ca.					
	DR	InterPro; IPR001438; EGF_III.					
	DR	InterPro; IPR001304; Lectin_C.					
	DR	InterPro; IPR000436; Sushi_SCR_CCP.					
	DR	Pfam; PF00008; EGF_2.					
	DR	Pfam; PF00059; lectin_C; 1.					
	DR	Pfam; PF00084; sushi; 1.					
	DR	PRINTS; PR00010; EGFBLOOD.					
	DR	SMART; SM00032; CCP; 1.					
	DR	SMART; SM00034; CLECT; 1.					
	DR	SMART; SM00179; EGF_Ca; 1.					
	DR	SMART; SM00001; EGF_like; 1.					
	DR	PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.					
	DR	PROSITE; PS00615; C_TYPELECTIN_1; 1.					
	DR	PROSITE; PS50041; C_TYPELECTIN_2; 1.					
	DR	PROSITE; PS00022; EGF_1; UNKNOWN_3.					
	DR	PROSITE; PS01186; EGF_2; 1.					
	DR	PROSITE; PS01187; EGF_Ca; 1.					
	KW	Calcium-binding; EGF-like domain; Glycoprotein; Repeat.					
	FT	NON_TER					
	SQ	SEQUENCE 315 AA; 36306 MW; 18132088897ED4AB CRC64;					
	Query Match	15.8%	Score 176; DB 4; Length 315;				
	Best Local Similarity	31.5%	Pred. No. 1e-08;				
	Matches	45;	Conservative 18; Mismatches 48; Indels 32; Gaps				
QY	17	QRPCYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIESEDEQKLIKEKIE	66				
	SEQUENCE 1479 AA; 166669 MW; 9F4BAF355F036FCE CRC64;						
	Query Match	15.8%	Score 176; DB 4; Length 1479;				
	Best Local Similarity	30.4%	Pred. No. 5.8e-08;				
	Matches	51;	Conservative 25; Mismatches 57; Indels 35; Gaps 7;				

RESULT 14

Q9XXX3 PRELIMINARY; PRT; 134 AA.

ID Q9XXX3; AC 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Receptor protein-tyrosine kinase (Fragment).
GN HTK28.
OS Hydra magnipapillata (Hydra);
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TAXID=6085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=105;
RX MEDLINE=20209407; PubMed=10744720;
RA Reidling J.C., Miller M.A., Steele R.E.;
RT "Sweet Tooth, a Novel Receptor Protein-tyrosine Kinase with C-type
Lectin-like Extracellular Domains";
RL J. Biol. Chem. 275:10323-10330(2000).
DR EMBL; AF129528; AAD30040.1; -.
DR HSSP; P22897; 1EGG.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPELECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPELECTIN_2; 1.
KW Kinase; Tyrosine-protein kinase.
FT NON_TER 1 1
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 15701 MW; E7B7211C881009BC CRC64;

Query Match 14.3%; Score 159.5; DB 11; Length 742;
Best Local Similarity 33.1%; Pred. No. 1.1e-06;
Matches 49; Conservative 17; Mismatches 55; Indels 27; Gaps 8;

QY 20 CYKVIYFHDTSRRLNFEAKEACRRDGGQLVSIESEDEQKLIKFENLLPSDGFHWIGL 79
Db 618 CY---YF---SLEKEILEDAKLFCEDKSSHLVFINSREEQQWIKKH---TVGRESHWIGL 668

QY 80 RRREEKQSNSTACQDLYAWTDGSISQFRNYYVDEP--CGSEEVVVMYHOPSAPAGIGGP 138
Db 669 TDSEQESE-----WKWLGDSPVDYKWNKAGQPDNWGSG----HGPGEDCA-GLIY 713

QY 139 MFQWNDDRCNMKNNFICKYSDEKPAVPS 166
Db 714 AGQWNDDFCDEINNFICE--KEREAVPS 739

Search completed: May 20, 2003, 17:06:04
Job time : 34 secs

RESULT 15

Q8VIF6 PRELIMINARY; PRT; 742 AA.

ID Q8VIF6; AC 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Scavenger receptor with C-type lectin.
GN COLEC12 OR SRCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TAXID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21575692; PubMed=11718900;
RA Nakamura K., Funakoshi H., Tokunaga F., Nakamura T.;
RT "Molecular cloning of a mouse scavenger receptor with C-type lectin
(SRCL)(1), a novel member of the scavenger receptor family.";
RL Biochim. Biophys. Acta 1522:53-58 (2001).
DR EMBL; AB038519; BAB82497.1; -.